VARIABLES AFFECTING THE LACTATION ESTIMATION CURVE IN WESTERN DAIRY CATTLE

445

by

VERN JAY CRANDALL

B. A., Brigham Young University, 1963

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements of the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1966

Approved by:

Stanley Wearden

TABLE OF CONTENTS

TABLE OF CONTENTS
[continued]

LIST OF TABLES AND FIGURES

TABLES

LIST OF TABLES AND FIGURES
[continued]

TABLES
[continued]

LIST OF TABLES AND FIGURES
[continued]

TABLES
[continued]

LIST OF TABLES AND FIGURES
[continued]

TABLES
[continued]

FIGURES

LIST OF TABLES AND FIGURES
[continued]

FIGURES
[continued]

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1. Motivation for the study.

An important phase of dairy herd management has to do with elimination of low producing cows. Because of impracticality, herds cannot begin with all cows being at exactly the same stage of lactation at the same time nor will they all be the same age. Many times it is also economically impractical to wait until all cows have completed a lactation before making culling decisions. For these reasons, it is necessary to find some way of converting the performance of all cows in a herd to a standard basis with respect to milk and butterfat production. An obvious procedure is to convert the production of each cow to expected production given the same environment, level of maturity, and a completed lactation of 305 days. One could then compare all cows on the same basis.

1.2. Literature concerning factors for converting records to 305 days.

1.2.1. Preliminaries.

Since converting cows to a standard basis for comparison is not a new problem, and since there are other applications of the extension of partial production records to 305 days such as sire proving, evaluating sold and dead cows, etc., tables of factors for converting milk and fat production to 305 days from a partial record have already been prepared, and the effects of certain environmental variables upon the shape of the lactation curve as expressed by these factors have been studied.

1.2.2. Types of factors. Environmental variables have been studied in terms of three types of factors. The first type consists of those for finding a quick estimate of the producing ability of a cow when full information is lacking. Such factors were determined from which a one-day test could be used to estimate producing ability. Adjustments were presented by Cannon *et al*. [1942], Madden *et al*. [1956, 1959], Lamb and McGilliard [1960a], and Van Vleck and Henderson [1961b]. The second type is used for estimating total production when several months records are available. Van Vleck and Henderson [1961b, e] presented a set of regression coefficients to be used when *m* sequential monthly records were available, either at the beginning or the end of the lactation. The third type consists of those based on cumulative production. Cumulative factors were presented by Cannon *et al*. [1942], Madden *et al*. [1956, 1959], Lamb and McGilliard [1960a], and Van Vleck and Henderson [1961b, c, e]. This third type is of most interest to the present study.

1.2.3. Variables studied. The cumulative factors mentioned above were studied in terms of breed, age, lactation number, milking frequency, level of production, herd, and season, though not all variables were examined in all the above studies. Of all variables considered, age was most important. Factors could be classified into two groups, i.e., those for cows less than 3 years old, and those for cows greater than or equal to 3 years of age. It was found that even when records were adjusted for age before determining the coefficients, all age

effects were still not removed, and thus, two corrections were needed, one to adjust for the effect of age on the shape of the lactation curve, and the other to adjust for differences in level of production between older and younger cows. [See Madden et al., 1955, 1959, and Lamb and McGilliard, 1960a.] Van Vleck and Henderson [1961b, c] also found it necessary to adjust for age. Little difference was found by Lamb and McGilliard [1960a] between the effect of age and that of lactation number.

Breed was found to be significant by Lamb and McGilliard [1960a], while Cannon et al. [1942] found that because of the similarity among regression coefficients for the various breeds studied, the coefficients could be combined. Madden et al. [1955, 1956, 1959] and Van Vleck and Henderson [1961b, c, e] studied only Holsteins.

Lamb and McGilliard [1960a] and Van Vleck and Henderson [1961b, c] found season to be significant, and Lamb and McGilliard found a breed by season interaction, with Holsteins being least influenced by season of freshening.

Madden et al. [1959] found milking frequency and level of production to be nonsignificant. Lamb and McGilliard [1960a] found herd differences to be non-significant, while Van Vleck and Henderson [1961b, e] presented their factors both on a within-herd basis and ignoring herd effects. In a later study, Van Vleck and Henderson [1961e] found that in estimating the regression ignoring herds [on a total regression rather than on an intra-herd regression basis] the inaccuracy

due to ignoring herd differences was greatly outweighed by the difficulty in obtaining the herd means and the inaccuracy of their estimates.

Most tables of estimates for extending records are based on either milk or fat alone but are used to predict both. It was found by Lamb and McGilliard [1960a] that there are definite differences between milk and fat adjustments; with the exception of the first month for Holsteins and Brown Swiss, cumulative factors for milk will underestimate butterfat production for the lactation. They also found that season of freshening was important for milk, but not for fat. Van Vleck and Henderson [1961b] found coefficients for milk and fat to be distinctly different. Madden et al. [1955] presented slightly different factors for milk and fat. Van Vleck and Henderson [1961c] raised the question whether milk and fat should both be extended, as this could change the per cent fat for the total lactation. The systems now used do not affect the fat percentage since they are based solely on milk or on fat.

1.2.4. Procedures Used. In the papers mentioned above which attempt to evaluate effects of certain variables on milk and fat production, two approaches are used: [1] a ratio factor is constructed based upon either the total milk produced to date divided into the total milk produced in 305 days, the total milk produced on $\underline{x}$ consecutive test days divided into the total milk produced on the 10 consecutive test days, or the sum of the reciprocals of the non-cumulative factors, the sum

then being reciprocated, for each x periods. [By non-cumulative is meant the per cent of the total milk produced each month of the lactation, or on each test day of the lactation.] This will give a cumulative ratio factor. [2] a regression coefficient, expressed either in terms of a regression equation or a single coefficient, is found from the regression of total milk produced in 305 days on total milk produced to date. In each of these methods the factor is adjusted in some way for such variables as age, season of freshening, parity, and breed.

The comparative accuracy of the above two approaches has been discussed by Madden et. al. [1959]. They found that the ratio method may under-estimate the total production of low producing cows and over-estimate total production of high producing cows since the ratio method adjusts only for the incompleteness of the lactation and does not take into account the incomplete repeatability of the parts of the lactation, which is a part of the regression method. This difference is largest during the early months of a lactation. The regression method, in which the cumulative part of the lactation is multiplied by the regression [b] of whole on part and added to an appropriate constant [a] to estimate the total production, not only adjusts for incompleteness, but also for unidentified sources of variation which make the part greater than or less than the average. The total estimated by regression varies less than the actual total. Although the variance of total production estimated by the ratio method is greater than that from the estimate found by regression, it has variation similar to

actual total production. For this reason, if these extended records are to be used for culling purposes or sire proving, ratio techniques are to be preferred. Since records extended by regression differ less than actual, this tendency to group the records around the mean may make selection decisions more difficult.

1.2.5. <u>Correlations of individual month records with complete record</u>. The part and the whole are always correlated; Van Vleck and Henderson [1961b] give the correlation of the individual month's record with the complete record as 0.57, 0.75, 0.81, 0.85, 0.85, 0.85, 0.83, 0.78, 0.66, and 0.53 for months 1,..., 10, respectively, on a within-herd basis. The correlations for the cumulative months with the complete record [on a within-herd basis] are 0.57, 0.75, 0.82, 0.87, 0.90, 0.93, 0.95, 0.97, and 0.99 for months 1,..., 9, respectively. When herd is not considered [see Van Vleck and Henderson (1961e)] the following correlations are obtained for the individual months 0.67, 0.82, 0.86, 0.89, 0.90, 0.89, 0.87, 0.81, 0.68, and 0.52 for months 1,..., 10, respectively. The correlations for the cumulative months records are 0.67, 0.82, 0.87, 0.91, 0.93, 0.95, 0.97, and 0.99 for months 1,..., 9, respectively.

1.3. <u>Literature concerning factors for converting records to mature equivalent</u>.

1.3.1. <u>Preliminaries</u>. If one desires to convert all cows in a herd to the same basis, when the cows are at different stages of lactation, by extending their production to 305 days and converting them to mature equivalent, one must not only be

concerned with extrapolation, but also with age correction adjustments.

Milk production increases with age at an ever-decreasing rate until maximum production is reached at 6 to 8 years; it then declines with advancing age. The regression of production on age is distinctly curvilinear. Lush and Shrode [1950] have given the theory and problems of age adjustments and their calculation. They described two methods and the bias associated with each. When the age corrections are calculated by Method A, averaging all records made at each age, concurrent selection introduces a bias. If the factors are calculated by Method B, comparing only records made by the same cows at two successive ages, concurrent selection introduces a bias in the opposite direction. The true age-change would then lie somewhere between the apparent changes given by the two methods.

Lush and Shrode also bring out the fact that a bias is introduced if the average inherent productivity of the dairy population is increasing. For instance, at any given date the averages of the older cows do not yet include the records from cows born in the most recent years when the production was actually higher. Henderson <u>et</u> <u>al</u>. [1959] presented techniques to estimate this genetic time trend with age.

1.3.2. <u>Studies of the effects of age on production</u>. The effects of age on production have been studied by workers other than Lush and Shrode [1950]. Searle and Henderson [1959, 1960], Searle [1960, 1961a] and Van Vleck and Henderson [1961c] have

presented various approaches and techniques to the calculation of adjustments for the effects of age on 305-day production.

1.3.3. Which factors should be used. Of all the adjustments in present use, the standard DHIA factors given by Kendrick [1955] are probably the most widely used. Miller's study [Miller, 1964] would indicate that in the absence of further research, these factors would be the wisest to use. Thus no attempt was made to estimate age-correction factors in the present study. Other reasons will be given in Section 1.5.

1.4. Literature concerning statistical considerations.

1.4.1. Preliminaries. In the statistical analyses of variables which affect the shape of the lactation curve, there are several conditions which cause problems with respect to the use of standard analysis of variance and regression techniques.

1.4.2. Non-orthogonal data. When dealing with analyses in which the numbers in each subclass are unequal, the problem becomes difficult, because standard analysis of variance techniques assume that degrees of freedom on which the variances for each of the effects being analyzed are based are orthogonal to [or independent of] each other. When unequal subclass frequencies are encountered, orthogonality cannot be guaranteed. Several papers have been written proposing either exact or approximate solutions to this problem. [See Yates, 1934; Snedecor, 1934; Harvey, 1960; Gosslee and Lucas, 1965; Wakefield, 1965; and Mielke and McHugh, 1965]. Several of these techniques are listed and discussed in various statistical books, for example, Snedecor [1956], Goulden [1952], Brownlee [1960],

Graybill [1961], and Dixon [1965].

Among techniques presented by Yates [1934] was a simple approximate method which consisted of ignoring differences between subclass numbers and performing an analysis of subclass means, assuming the variances of the means to be equal and considering the means as one observation per cell. The analysis is then performed using techniques for equal subclasses. This approximation is only useful, he pointed out, if the subclass numbers do not differ greatly. Recently, Gosslee and Lucas [1965] concluded, concerning the question of level of significance in the method of unweighted squares of means, that the effects of unequal variance among the cell means have only a moderate effect.

Estimates of components of variance are also affected by unequal subclass numbers. Methods of dealing with this problem have been presented by Henderson [1953], Harvey [1960], and Bush and Anderson [1963]. However in the case of unweighted squares of means [Yates, 1934, and Gosslee and Lucas, 1965], it is necessary only to solve the expected mean squares in the usual manner for orthogonal data.

1.4.3. Multiple comparisons. Use was made of Duncan's Multiple Range Test to test for differences among means in several of the papers discussed in Sections 1.2 and 1.4.2. Although Duncan's test decreases the probability of a Type II error, the probability of a Type I error inflates until no estimate of it even exists. For this reason, tests for which the probability of the Type I error remains constant

and known, such as Tukey's hsd or Sheffé's test, are to be preferred, [see Federer, 1955, Brownlee, 1960, and Sheffé, 1959].

1.4.4. <u>The comparison of growth curves</u>. Comparing several lactation curves is similar to the problem of comparing growth curves. Rao [1958] defines this problem in terms of comparing the characteristics of growth under different conditions such as diet, environment, etc. The difficulty in comparing growth curves based on values at a number of points along the curve can be overcome by reducing the data to the lowest possible dimensions without sacrificing the essential information. This can lead to a more efficient procedure. As an example of this, he cites the fitting of a second degree polynomial to each growth curve classified by the factors to be studied. The coefficients of the linear and quadratic terms can then be taken to represent the salient features of the growth, and thus a large number of observations can be replaced by two coefficients. A small number of parameters to use in the determination of differences between groups of growth curves can thus be found.

1.4.5. <u>The distribution of the ratio of two normally distributed random variables</u>. The lactation estimation curve is generally expressed as a series of 10 ratios between the total milk produced to date and the total 305-day milk production, where these two values become more and more highly correlated as the ratios are taken from the first to the tenth period, being exactly 1.00 for the tenth period. These ratios are then ratios of two random variables which are assumed to be normal. The distribution of the ratio of two normally distributed random

variables is <u>not</u> normal. Geary [1930], Fieller [1932], Merrill [1928], and Marsaglia [1965] have considered this problem in general, and in its specific applications, such as bioassay and sampling theory, it has been discussed by Fieller [1940], Finney [1964], Hansen, *et al*. [1956], and Deming [1950, 1960].

Although the distribution of $z = y/x$ is not normal, there exists a function of z which is approximately normal. If x, y are distributed normally with correlation r and with means 0 and standard deviations α and β, respectively, then if $z = \frac{b + y}{a + x}$, where a and b are non-negative constants, then $t = \frac{az - b}{\sqrt{\alpha^2 z^2 + 2r\alpha\beta z + \beta^2}}$ is distributed $N[0,1]$ if $a + x$ is unlikely to assume negative values. This condition is amply satisfied if $a \geq 3\alpha$, i.e., if the coefficient of variation of $a + x$ is not greater than 1/3. Thus, normal theory can be used when ratios make up the observations and the x and y values are positive and not close to zero and are large with respect to their standard deviations, provided the above transformation is made.

1.5. <u>The purpose of this thesis</u>. The purpose of this thesis is to examine the effects of certain variables upon the shape of the lactation estimation curve, using a different approach to the problem than those which have been heretofore presented, taking into account the findings presented in the various sections of this introduction. The data, from herds in the Western states--data which have not previously been available, will be used to provide factors based upon the significant

variables for use in estimating total from part lactation production for both milk and fat.

Because of the nature of the data available, it was not practical to construct factors for converting cow records to mature equivalent. For this reason, and because of the availability of the age correction factors already in existence, it was decided not to consider them in the present study.

CHAPTER 2

METHODS AND MATERIALS

2.1. Materials. Data used in this study came from the files of the DHI Computing Service, Provo, Utah, and comprised all records tabulated by the computing service from November, 1962, through March, 1965, excluding records for the state of California. 1,400,800 monthly cow production totals were recorded on magnetic tape and sorted into sequence by month of production within cow. The records were then edited to satisfy the following conditions.

1. A lactation consisting of 305 to 350 days. [Note that when a lactation exceeded 305 days, the production was converted back to 305 days.] There is some question whether to use short lactation records, i.e., completed lactations of cows which have been in milk more than, say, 270 days but less than 305, because a large number of cows complete their lactations before reaching 305 days.

2. No missing test days during the lactation. Sometimes, if a cow is purchased fresh, the first part of her lactation will be missing. Although it was highly unlikely because of the way in which the records were assembled, test dates were tested for consecutiveness to eliminate lactations with non-consecutive dates.

3. The first month's test being used to estimate less than or equal to 50 days' production. This restriction was used to avoid overlapping of effects from the first and second period's production, as this could mask or confound the effects

of number of days into the lactation on the lactation estimation curve. At the DHI Computing Service, as soon as a cow exceeds 320 days in milk without completing her lactation, i.e. approximately one month after reaching 305 days, the lactation is automatically flagged and a 305 day record computed. Because of the choice of 50 days as the cut-off for the first period, 350 days becomes the maximum possible number of days in the flagged lactation. As was mentioned before, such records were then converted back to 305 days.

4. The latest record being used if duplicate records appeared for any given month. It was felt that if such duplicates appeared, the latest record would be the most likely to be correct.

5. Two-time milking. There were not enough three times per day milking records to make their consideration worthwhile.

6. A lactation number reported. In these records, the age of the cow was not available. Since age is one of the variables to be studied [see sections 1.2.2, and 2.2.1], a record was worthless unless some estimate of age was available, thus lactation number was necessary.

7. No correction appearing anywhere in the lactation. Often in the reporting of information, mistakes and inconsistencies occur. When these mistakes are found, a correction card is placed in the file to adjust the record for the cumulative production. A correction card is also used to estimate missing production, correct for sickness, etc. Because these corrections indicate incorrect data in the early part of the lactation which could affect the shape of the lactation estimation curve, records containing them were eliminated.

8. Consistency of data. Whenever data were obviously incorrect or inconsistent, i.e., milk weights that were too large [99999 in that portion of the record, etc.], or when data needed for the analysis were missing, such lactations were eliminated.

It should be noted with respect to these data that not only was the age of the cow not available, but also, it was not possible to determine breed from the data.

For the analysis [which will be referred to later] of Days Dry and Days Open, a further edit was performed. In addition to the above conditions, those data used in this analysis were also required to have a non-zero number of days carried calf and be second lactations or greater.

There were 18,541 records satisfying the above 8 conditions, and 13,023 satisfying also the conditions mentioned in the preceding paragraph.

Examples of the form of the data before and after editing appear in Appendix I.

2.2. Methods.

2.2.1. Variables to be studied.

It was decided to compare lactation estimation curves based on the edited data with respect to the following variables.

1. Geographic Area. It was decided to examine the effects of different groupings of the states represented in these data.

2. Average per cent fat. Because breed was not available to be used as a variable in this study, average per cent fat was substituted. Average per cent fat is defined as total

pounds of milk divided into total pounds of fat, giving the weighted average over the entire lactation. An interesting consideration of the use of this variable rather than breed comes from the question of what the real breed differences are in terms of milk and fat production. For example, what is the difference between a Guernsey that tests 3.65% and a Holstein that tests 3.65% fat? Or a Holstein that tests 4.75% and a Guernsey that tests 4.75% fat?

3. Lactation number. In all studies referred to in Section 1.2 in which they were considered, both age and lactation number were found to be significant. Lamb and McGilliard [1959] found that except in cases where age and lactation number did not coincide, lactation number was actually a better means of estimation.

4. Season of Freshening. In several of the studies discussed in Section 1.2, season was found to have an effect on the shape of the lactation estimation curve.

5. Previous days dry. As Smith [1959] and Johansson [1961] point out, previous days dry ont only have an effect on the current lactation, but also on the succeeding one.

6. Days open. Days open is defined to be the difference between days carried calf and days in milk plus three days. In a discussion of physiology of lactation, both Smith [1959] and Johansson [1961] mentioned that this variable was found to have an effect on the lactation.

7. Level of milk production. In all factors in use at the present time, level of milk production is not taken into account. Obviously, it is impossible to consider factors which take this into account in terms of total milk production

for a given lactation, unless one uses level of production for the preceding lactation, level of production for the first month, or the herd average as an indication of level of production of the current one.

8. Level of fat production. It was decided to examine also the level of fat production to see if the make-up of the production cycle was affected by differences in level of production. It is known that the per cent fat test decreases as the milk production increases, but not what happens to the shape of the curve as milk or fat production increases.

2.2.2. Levels and coding of the variables to be used. To determine the levels of the above factors to be used, several things were taken into account, the most important being the frequencies of numbers of cows in each grouping and the physiological and environmental effects involved.

The number of records from each state are given in the following table.

Table 1

The number of records from each state used in the study.

State	Number of Records
North Dakota (42)	19
Iowa (45)	28
Nebraska (47)	467
Montana (81)	718
Idaho (82)	553
Wyoming (83)	775
Colorado (84)	4,715
New Mexico (85)	1,154
Arizona (86)	1,440
Utah (87)	8,210
Nevada (88)	503
	18,582

When the analyses were begun, certain data were found to be incorrect and had not been detected in the original edit. When these were eliminated, the number of records was reduced from 18,582 to 18,541. The states were combined such that North Dakota, Iowa, Nebraska, Montana, Idaho, and Wyoming were considered Area 2; Colorado was considered Area 3; New Mexico and Arizona were considered Area 4; and Utah and Nevada were considered Area 5. These groupings were made on the basis of the above frequency table [Table 1], climatic similarities, and the necessary combinations to eliminate missing cells when the other factors in the analysis were taken into account.

The levels of average per cent fat were determined from an examination of the average per cent fat for the different breeds, and also from the mean and variance of the frequency distribution for Utah. The mean was 3.8% and the standard deviation was 0.6%. When this was done, the levels were considered to be those shown in Table 2.

Table 2

The data classes of per cent fat and the codes thereof [preliminary].

Per Cent Fat	Code
2.8% and under	1
2.9% to 3.4%	2
3.5% to 4.0%	3
4.1% to 4.6%	4
4.7% to 5.2%	5
5.3% to 5.8%	6
5.9% and over	7

When these levels were examined for cell frequencies, it was found necessary to make the following combinations and to

analyze the data from Geographic Areas 3 and 5 [as presented in Tables 3 and 4].

Table 3

The data classes of per cent fat and the codes thereof for Geographic Areas 2 and 4.

Per Cent Fat	Code
3.4% and under	2
3.5% to 4.0%	3
4.1% and over	4

Table 4

The data classes of per cent fat and the codes thereof for Geographic Areas 3 and 5.

Per Cent Fat	Code
3.4% and under	2
3.5% to 4.0%	3
4.1% to 4.6%	4
4.7% to 5.2%	5
5.3% and over	6

These levels correspond roughly to low fat producing Holsteins [2], mostly Holsteins and Milking Shorthorns with some Ayrshires and Brown Swiss [3], and a few Holsteins and Milking Shorthorns, the remainder being Ayrshires, Brown Swiss, Guernseys, and Jerseys [4] for Geographic Areas 2 and 4, and to low fat producing Holsteins [2], mostly Holsteins and Milking Shorthorns, with some Ayrshires and Brown Swiss [3], half of the Ayrshires and Brown Swiss, and half of the Guernseys plus a few Holsteins and Milking Shorthorns [4], Guernseys and half Jerseys [5], and predominantly Jerseys with a few Guernseys [6] for Geographic Areas 3 and 5.

Although there has been strong evidence that, in general,

age can be divided on "less than 3 years" and "3 years and over" [see Section 1.2], it was decided to look at age [based on lactation number] in terms of the six levels given in Table 5.

Table 5

The data classes of lactation number and the codes thereof.

Lactation Number	Code
First	1
Second	2
Third	3
Fourth	4
Fifth	5
Sixth and over	6

It was decided to group season of freshening in the manner indicated in the following table.

Table 6

The data classes of season and the codes thereof.

Season	Code
December, January and February	1
March, April, and May	2
June, July, and August	3
September, October and November	4

Previous days dry were divided into six groups. Johansson [1961] stated that the optimum dry period was from 35-40 days for the Swedish-Friesian breed, while Smith [1959] found it to

be about 55 days for a constant calving interval of 365 days for 10,000 pounds of milk in American breeds. An examination of the frequency distribution for previous days dry showed that the largest numbers of records were for the period from 47 to 68 days with the largest number being for 59 days. The number of days previously dry ranged from one day to over 328 days, with the numbers increasing to 59 days and decreasing thereafter. [This discussion pertains to Utah cows, as there were more of them than those for any other state. The frequency distributions were considered on a within-state basis.] After studying the frequency distribution, it was decided to subdivide days previous dry in the manner indicated in Table 7.

Table 7

The data classes of previous days dry and the codes thereof.

Previous Days Dry	Code
20 days and under	1
21 days to 40 days	2
41 days to 60 days	3
61 days to 80 days	4
81 days to 100 days	5
100 days and over	6

Note that previous days dry cannot be considered for first lactations, and therefore they were dropped from this part of the analysis.

A study of the frequency distribution for Utah for days open showed that the number of cows per day was quite small up to 69 days open. From 69 days open to 200 days, the number of cows per day is quite uniform, although it reaches a maximum at 92 days and is somewhat lower at the ends of this

interval. This variable was arbitrarily subdivided into the intervals described in the table below.

Table 8

The data classes of days open and the codes thereof.

Days Open	Code
60 days and under	1
61 days to 80 days	2
81 days to 100 days	3
101 days to 120 days	4
121 days to 140 days	5
141 days to 160 days	6
161 days to 180 days	7
181 days and over	8

The following data classes for level of milk production were determined by examining the frequency distribution for Utah.

Table 9

The data classes of level of milk production and the codes thereof.

Level of Milk Production	Code
5,999 pounds and less	1
6,000 pounds to 6,999 pounds	2
7,000 pounds to 7,999 pounds	3
8,000 pounds to 8,999 pounds	4
9,000 pounds to 9,999 pounds	5
10,000 pounds to 10,999 pounds	6
11,000 pounds to 11,999 pounds	7
12,000 pounds to 12,999 pounds	8
13,000 pounds to 13,999 pounds	9
14,000 pounds to 14,999 pounds	10
15,000 pounds to 15,999 pounds	11
16,000 pounds to 16,999 pounds	12
17,000 pounds to 17,999 pounds	13
18,000 pounds to 18,999 pounds	14
19,000 pounds to 19,999 pounds	15
20,000 pounds and greater	16

Level of fat production was divided into classes by examining the mean and variance of the frequency distribution

for Utah. The mean was found to be approximately 460 pounds of fat, with a standard deviation of approximately 100 pounds. An interval of roughly 1/3 the standard deviation was used and the following levels were determined.

Table 10

The data classes of level of fat production and the codes thereof.

Level of Fat Production	Code
309 pounds and less	1
310 pounds to 339 pounds	2
340 pounds to 369 pounds	3
370 pounds to 399 pounds	4
400 pounds to 429 pounds	5
430 pounds to 459 pounds	6
460 pounds to 489 pounds	7
490 pounds to 519 pounds	8
520 pounds to 549 pounds	9
550 pounds to 579 pounds	10
580 pounds to 609 pounds	11
610 pounds to 639 pounds	12
640 pounds to 669 pounds	13
670 pounds to 699 pounds	14
700 pounds and greater	15

2.2.3. Classification of the analyses to be performed. As can be seen, if these eight variables were analyzed simultaneously, one would need several million cells. For this reason, the over-all analysis was broken into the following analyses.

1. A four-way analysis of Geographic Area, Average Per Cent Fat, Lactation Number, and Season for Geographic Areas 2 and 4.

2. A four-way analysis of Geographic Area, Average Per Cent Fat, Lactation number, and Season for Geographic Areas 3 and 5.

3. A two-way analysis of Days Previous Dray and Days Open.

4. A one-way analysis of Level of Milk Production.

5. A one-way analysis of Level of Fat Production.

Two sets of these analyses were performed; one for milk production, and one for fat production.

2.2.4. <u>Nature of the statistical problems involved and their solutions</u>.

2.2.4.1. Before presenting the models and the analyses for the five classifications of Section 2.2.3, a discussion of the problems involved seems necessary. To justify the solutions obtained for some of these problems, the discussions of statistical considerations of Section 1.4 will be referred to.

2.2.4.2. <u>The ten period approach vs. the 305 day approach</u>. There are two ways of expressing a lactation curve. The first is to present the ten monthly test days, or the monthly production. The problem then consists of studying ten equally spaced observations for each cow. The second method is more closely associated with the actual Dairy Herd Improvement Association rules and practices. Each month's production is based on one test for that month, and the results are multiplied by the number of days in the month. To make this test meaningful, because of the gradual decline in production over the lactation, a centering date is introduced, causing the production to be estimated for 15 days prior to that date and 12 to 15 days after that date, depending on the month of test. The test date should fall within 3 days of the centering date, but in practice, the difference is sometimes greater. Because a cow cannot be

analysis, where "X = days in milk" can take on the values 1,..., 305. The first problem is that there will not be the same number of observations for each record. The second is that the X values will not be the same for all records. Despite the fact that x, y and N are random, once they are determined, X can be considered a set of fixed values.

It is difficult to use standard statistical analyses of the 8 to 11 observations in comparing the lactation estimation curves. Thus, it was decided to analyze differences between the lactation estimation curves by using the regression of Y on X, where Y is the ratio of the cumulative monthly production to the total production in 305 days. This reduces the comparison of a possible 11 observations per record to one observation per record. Fitting orthogonal polynomials to factors representing the shape of the lactation estimation curve indicated that when this approach was used, it was necessary to use a fourth degree polynomial to adequately represent the curve. This would give four coefficients to analyze instead of one. [Crandall, 1963]. It would be better to get the curve in terms of one coefficient. [See Rao [1958] and Section 1.5.3].

2.2.4.4. <u>The problem of the distribution of a ratio and its solution</u>. Since the points on the lactation estimation curve are defined to be

$$Y_X = \frac{\text{Total 305-day production}}{\text{Total production in } X \text{ days}} ,$$

and since both the numerator and the denominator are assumed to be normally distributed, it is necessary to determine the

distribution of Y. From the articles by Geary [1930], Fieller [1932], and Marsaglia [1965] cited in Section 1.4.5, since both the numerator and the denominator are positive and not close to zero, the coefficients of variation being small, the distribution of a function of Y can be found which is approximately normal. Because of the difficulty in calculating the necessary standard deviation and correlations, the data were not transformed to this function which would normalize the observations. Instead it was hoped that the use of the transformation of Section 2.2.4.5 and the use of means of observations [Section 2.2.4.6] would "smooth" the data and cause them to approach a normal distribution.

2.2.4.5. <u>The problem of a transformation and its solution</u>. The lactation estimation curve is distinctly curvilinear, and thus if linear regression, as discussed above [Section 2.2.4.3] were used, a poor fit would result. Instead the curvilinear model was assumed to be

$$Y = A[305/X]^C. \qquad [5]$$

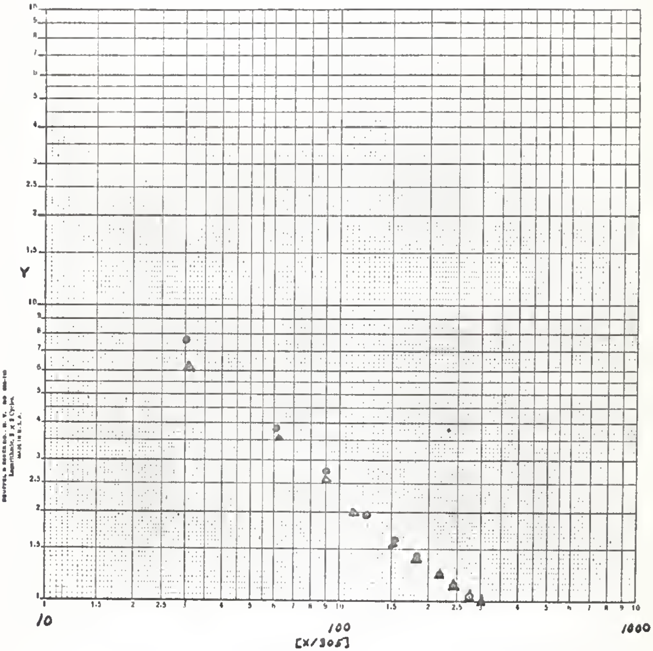
In order to fit this equation, the log of both sides was taken to give,

$$\log_{10} Y = \log_{10} A + C \log_{10} [305/X]. \qquad [6]$$

This equation was fitted for each record using standard regression techniques [see Appendix I.] The analysis of variance for regression on a small portion of the data indicated extremely high significance. When two cases of

FIGURE 1

TWO FITTED REGRESSION LINES [LOG-LOG TRANSFORMATION]



$Y = A[X/305]^C$

actual data were plotted on log-log paper, [Figure 1] there was still a slight curvelinear effect after the transformation. However, when overall factors obtained by this method were compared with factors in current use, it was found that there was a very close fit.

Figure 1 illustrates that heterogeneity among a group of fitted regression lines will cause the largest differences to occur early in the lactation, and the lines to converge to one at 305 days. [Note the difference between Equation (5) and that of Figure 1. The regression coefficients were all calculated using the equation in Figure 1, which yielded a negative coefficient for all records. By omitting the minus sign, for all coefficients, one has the same results as would have been obtained had Equation (5) been used.] The reason why this heterogeneity exists is because Y_X does not have uniform variance for all values of X. When X is small, the variance of Y_X is extremely large, and as X approaches 305 days, the variance of Y_X approaches zero. This effect of the variance violates one of the assumptions of covariance analysis, i.e., that of homogeneous variance along the regression line, but does not affect the normality of the regression coefficients. Thus analyses of variance can be performed using the regression coefficients as observations, but analysis of covariance techniques cannot be used.

When the regression coefficients were examined, it was found that the values of A [Equation (5)] were practically one. thus this term was not considered in later studies. Actually,

a more accurate model would have been obtained by ignoring the mean and recalculating the regression coefficients, thus giving

$$Y = [305/X]^C. \qquad [7]$$

In this case, $C = \Sigma X'Y'/\Sigma X'^2$, where $X' = \log [305/X]$ and $Y' = \log Y$, instead of $C = \Sigma x'y'/\Sigma x'^2$, where $x' = [X' - \overline{x}']$ and $y' = [Y' - \overline{y}']$. While this difference is unimportant with respect to the regression coefficient, C, it does make a difference if covariance analysis is to be used. The fact that one must subtract $[\Sigma X']^2/N$ from the denominator of the standard error makes a larger standard error when the mean is considered, incorrectly, in the model, than if the mean were not considered. For these reasons, it was decided to perform analysis of variance techniques with the regression coefficients for each record as the observations, rather than applying covariance techniques.

2.2.4.6. The problem of unequal subclass frequencies and its solution. Because of unequal subclasses involved in this study, least squares techniques seemed applicable, however, none of the least squares computer programs available at either Brigham Young University [IBM 7040] or the University of Washington [IBM 7094] would handle as large an analysis as the first two described in Section 2.2.3. In looking for an alternate method, it was decided to use the method of unweighted means [Yates, 1934]. A second consideration for this particular analysis was that it was felt that the means of the regression coefficients would be more nearly normal. For this method, the means were analyzed according to a standard orthogonal

factorial design for the first three analyses described in Section 2.2.3 and a one-way analysis of variance with unequal subclass frequencies was used for the last two analyses described. In the factorial analyses, the highest order interaction was used as an estimate for error.

2.2.5. <u>The linear statistical models for the five analyses</u>. All effects are considered fixed, with the exception of the error term.

For the first analysis for areas 2 and 4, the regression coefficients for milk and fat were each assumed to be described by the following linear model

$$y_{ijkl} = \mu + a_i + b_j + c_k + d_l + ab_{ij} + ac_{ik} + ad_{il} + bc_{jk} + $$
$$+ bd_{jl} + cd_{kl} + abc_{ijk} + abd_{ijl} + acd_{ikl} + bcd_{jkl} + e_{ijkl};$$

$i = 2,4;$ $\quad k = 1,\ldots, 6;$

$j = 2,\ldots, 4;$ $l = 1,\ldots, 4;$ for all y_{ijkl} in areas 2 and 4.

where

y_{ijkl} = the regression coefficient in the i^{th} geographic area, in the j^{th} per cent fat level, in the k^{th} lactation, and in the l^{th} season

μ = the overall regression coefficient for the population, when equal frequencies exist in all subclasses.

a_i = the effect of the i^{th} geographic area.

b_j = the effect of the j^{th} per cent fat level.

c_k = the effect of the k^{th} lactation.

d_l = the effect of the l^{th} season.

ab_{ij} = the effect of the ij^{th} subclass of geographic area and

per cent fat after the average effects of the i^{th} geographic area and the j^{th} per cent fat level have been removed.

ac_{ik} = the effect of the ik^{th} subclass of geographic area and lactation number after the average effects of the i^{th} geographic area and the k^{th} lactation have been removed.

ad_{il} = the effect of the il^{th} subclass of geographic area and season after the average effects of the i^{th} geographic area and the l^{th} season have been removed.

bc_{jk} = the effect of the jk^{th} subclass of per cent fat and lactation after the average effects of the j^{th} per cent fat level and the k^{th} lactation have been removed.

bd_{jl} = the effect of the jl^{th} subclass of percent fat and season after the average effects of the j^{th} per cent fat level and the l^{th} season have been removed.

cd_{kl} = the effect of the kl^{th} subclass of lactation and season after the average effects of the k^{th} lactation and the l^{th} season have been removed.

abc_{ijk} = the effect of the ijk^{th} subclass of geographic area, per cent fat, and lactation number after the average effects of the i^{th} geographic area, the j^{th} per cent fat level, the k^{th} lactation, the ij^{th} subclass of geographic area and per cent fat, the ik^{th} subclass of geographic area and lactation number, and the jk^{th} subclass of per cent fat and lactation number have been removed.

abd_{ijl} = the effect of the ijl^{th} subclass of geographic area, per cent fat, and season after the average effects of the i^{th} geographic area, the j^{th} per cent fat level, the l^{th} season, the ij^{th} subclass of geographic area and per cent fat, the il^{th} subclass of geographic area and season, and the jl^{th} subclass of per cent fat and season have been removed.

acd_{ikl} = the effect of the ikl^{th} subclass of geographic area, lactation number, and season after the average effects of the i^{th} geographic area, the k^{th} lactation number, the l^{th} season, the ik^{th} subclass of geographic area and lactation number, the ik^{th} subclass of geographic area and season, and the kl^{th} subclass of lactation number and season have been removed.

bcd_{jkl} = the effect of the jkl^{th} subclass of per cent fat, lactation number, and season after the average effects of the j^{th} per cent fat level, the k^{th} lactation, the l^{th} season, the jk^{th} subclass of per cent fat level and lactation number, the jl^{th} subclass of per cent fat level and season, and the kl^{th} subclass of lactation number and season are removed.

e_{ijkl} = the residual error, the amount of variation not accounted for by the above effects, [the highest order interaction assumed to be zero]. These errors are assumed to be $NID[0, \sigma_e^2]$.

The various effects are expressed as deviations from the mean, and therefore, $\Sigma a_i = \Sigma b_j = \Sigma c_k = \Sigma d_l = 0$, and $\sum_i ab_{ij} = \sum_j ab_{ij} = \ldots = \sum_k cd_{kl} = \sum_l cd_{kl} = 0$. Similar restrictions are placed on the three-way interactions. Because all the effects in this study are assumed fixed, estimation of the constants for the levels of the significant effects were made to be used in estimating total production from part lactation records.

For the second analysis, for areas 3 and 5, the model is exactly the same, except that the subscripts, $i, \ldots, l$, are over different values. $i = 3, 5$; $j = 2, \ldots, 6$; $k = 1, \ldots, 6$; and $l = 1, \ldots, 4$; for all y_{ijkl} in areas 3 and 5. The above assumptions also hold in this model.

For the third analysis, for days dry and days open, the regression coefficients for milk and fat were each assumed to be described by the following model

$$y_{ijk} = \mu + a_i + b_j + e_{ij}$$

$$i = 1, \ldots, 6 \quad \text{and} \quad j = 1, \ldots, 8$$

where

y_{ij} = the regression coefficient in the i^{th} class of days previous dry and the j^{th} class of days open.

μ = the overall regression coefficient for the population when equal frequencies exist in all subclasses.

a_i = the effect of the i^{th} class of days previous dry.

b_j = the effect of the j^{th} class of days open.

e_{ij} = random errors not accounted for by the other terms in the model.

Since the interaction is assumed to be zero, the error term is represented by the two-way interaction, and the errors expressed thereby are assumed NID $[0, \sigma_e^2]$. It should be noted that the y_{ij} are from the entire population of areas 2 and 4, as well as 3 and 5, excluding the y_{ij} for which days carried calf equal zero and those for which the record is a first lactation.

For the fourth analysis, for level of milk production, the regression coefficients for milk and fat were assumed to be expressed by the following linear model

$$y_{ij} = \mu + a_i + e_{ij}$$

$$i = 1, \ldots, 16 \quad \text{and} \quad j = 1, \ldots, n_i$$

where

y_{ij} = the j^{th} regression coefficient in the i^{th} level.

μ = the overall regression coefficient for the population when equal frequencies exist in all subclasses.

a_i = the effect of the i^{th} level of milk production.

e_{ij} = random errors, which are NID$[0, \sigma_e^2]$.

Since the a_i are expressed as deviations about the mean, $\Sigma a_i = 0$. It should be noted that the y_{ij} are here from the entire population, areas 2, 3, 4, and 5.

For the fifth analysis, for level of fat production, the regression coefficients for milk and fat were assumed to be described by the following model

$$y_{ij} = \mu + a_i + e_{ij}$$

$$i = 1, \ldots, 15 \quad \text{and} \quad j = 1, \ldots, n_i$$

where

y_{ij} = the j^{th} regression coefficient in the i^{th} level.

μ = the overall regression coefficient for the population when equal frequencies exist in all subclasses.

a_i = the effect of the i^{th} level of fat production.

e_{ij} = random errors, which are NID$[0, \sigma_e^2]$.

The same comments hold here as for the fourth analysis above.

CHAPTER 3

RESULTS AND DISCUSSION

3.1. Results of the analyses of variance and discussion. The results of analyses 1 through 5 for milk and fat will be presented in the following tables.

Table 11

Analysis of regression coefficients for factors for estimating total milk production from cumulative milk production for geographic areas 2 and 4.

Source	df	Sums of Squares	Mean Squares	F
Geographic Area	1	0.00000	0.00000	0.00000
Per Cent Fat	2	0.01489	0.00745	17.32558**
Lactation Number	5	0.09036	0.01807	42.02326**
Season	3	0.03269	0.01090	25.34884**
GA x PF	2	0.00051	0.00026	0.60465
GA x LN	5	0.00561	0.00112	2.60465*
GA x S	3	0.00098	0.00033	0.76744
PF x LN	10	0.00388	0.00039	0.90698
PF x S	6	0.01487	0.00248	5.76744**
LN x S	15	0.01056	0.00070	1.62791
GA x PF x LN	10	0.00645	0.00064	1.48837
GA x PF x S	6	0.00174	0.00029	0.67442
GA x LN x S	15	0.00765	0.00051	1.18605
PF x LN x S	30	0.01984	0.00066	1.53488
Residual	30	0.01292	0.00043	
Total	143	0.22296		

Note that wherever [*] is used in an F-column, it means "significant at the α = 0.05 level", and wherever [**] is used in an F-column, it means "significant at the α = 0.01 level.

Table 12

Analysis of regression coefficients for factors for estimating total milk production from cumulative milk production for geographic areas 3 and 5.

Source	df	Sums of Squares	Mean Squares	F
Geographic Area	1	0.00113	0.00113	3.53125
Per Cent Fat	4	0.05572	0.01393	43.53125**
Lactation Number	5	0.11283	0.02257	70.53125**
Season	3	0.07059	0.02353	73.53125**
GA x PF	4	0.00412	0.00103	3.21875*
GA x LN	5	0.00228	0.00046	1.43750
GA x S	3	0.00021	0.00007	0.21875
PF x LN	20	0.01383	0.00069	2.15625*
PF x S	12	0.01114	0.00093	2.90625**
LN x S	15	0.00160	0.00011	0.34375
SA x PF x LN	20	0.00499	0.00025	0.78125
SA x PF x S	12	0.00285	0.00024	0.75000
SA x LN x S	15	0.00853	0.00057	1.78125
PF x LN x S	60	0.01413	0.00024	0.75000
Residual	60	0.01916	0.00032	
Total	239	0.32311		

Table 13

Analysis of regression coefficients for factors for estimating total milk production from cumulative milk production for all areas, excluding first lactation records and those for which there are no days carried calf.

Source	df	Sums of Squares	Mean Squares	F
Days Previous Dry	5	0.18584	0.03717	1.31204
Days Open	7	0.18667	0.02667	0.94140
Residual	35	0.99163	0.02833	
Total	47	1.36415		

Table 14

Analysis of regression coefficients for factors for estimating total milk production from cumulative milk production for all records in all areas.

Source	df	Sums of Squares	Mean Squares	F
Level of Milk Prod.	15	1.95564	0.13038	20.40378**
Residual	18,529	118.46874	0.00639	
Total	18,544*	120.42438		

*Note that there are 18,541 records used in the study, while 18,545 records were used in this analysis. This is because the sort routine used to set up this analysis filled in the four remaining positions in the last block of tape with four valid records from the last cell. Since there were 175 records in the last cell, it was felt that the effects of this duplication would be negligible.

Table 15

Analysis of regression coefficients for factors for estimating total milk production from cumulative milk production for all records in all areas.

Source	df	Sums of Squares	Mean Squares	F
Level of Fat Prod.	14	0.52779	0.03770	5.86314**
Residual	18,530	119.14661	0.00643	
Total	18,544*	119.67440		

*Note that the discussion following Table 14 applies here also, since there were 227 records in the last cell for this analysis.

Table 16

Analysis of regression coefficients for factors for estimating total fat production from cumulative fat production for geographic areas 2 and 4.

Source	df	Sums of Squares	Mean Squares	F
Geographic Area	1	0.00263	0.00263	5.49717*
Per Cent Fat	2	0.00607	0.00304	6.33333**
Lactation Number	5	0.13302	0.02660	55.41667**
Season	3	0.00663	0.00221	4.60417**
GA x PF	2	0.00286	0.00143	2.97917
GA x LN	5	0.00418	0.00084	1.75000
GA x S	3	0.00212	0.00071	1.47917
PF x LN	10	0.01006	0.00101	2.10417
PF x S	6	0.01252	0.00209	4.35417**
LN x S	15	0.00812	0.00054	1.12500
GA x PF x LN	10	0.00810	0.00081	1.68750
GA x PF x S	6	0.00402	0.00067	1.39853
GA x LN x S	15	0.00783	0.00052	1.08333
PF x LN x S	30	0.02747	0.00092	1.91666*
Residual	30	0.01455	0.00048	
Total	143	0.25018		

Table 17

Analysis of regression coefficients for factors for estimating total fat production from cumulative fat production for geographic areas 3 and 5.

Source	df	Sums of Squares	Mean Squares	F
Geographic Area	1	0.00094	0.00094	1.84314
Per Cent Fat	4	0.00874	0.00219	4.29412**
Lactation Number	5	0.20104	0.04021	78.84314**
Season	3	0.00476	0.00159	3.11765*
GA x PF	4	0.01281	0.00320	6.27451**
GA x LN	5	0.00206	0.00041	0.80392
GA x S	3	0.00031	0.00010	0.19608
PF x LN	20	0.00883	0.00044	0.86275
PF x S	12	0.00457	0.00038	0.74510
LN x S	15	0.00613	0.00041	0.80392
GA x PF x LN	20	0.00717	0.00036	0.70588
GA x PF x S	12	0.00363	0.00030	0.58824
GA x LN x S	15	0.00951	0.00063	1.23529
PF x LN x S	60	0.01801	0.00030	0.58824
Residual	60	0.03042	0.00051	
Total	239	0.31893		

Table 18

Analysis of regression coefficients for factors for estimating total fat production from cumulative fat production for all areas, excluding first lactation records and those for which there are no days carried calf.

Source	df	Sums of Squares	Mean Squares	F
Days Previous Dry	5	0.14375	0.02875	1.00067
Days Open	7	0.18927	0.02704	0.94678
Residual	35	0.99977	0.2856	
Total	47	1.33279		

Table 19

Analysis of regression coefficients for factors for estimating total fat production from cumulative fat production for all records in all areas.

Source	df	Sums of Squares	Mean Squares	F
Level of Milk Prod.	15	1.63407	0.10894	13.70314**
Residual	18,529	147.30780	0.00795	
Total	18,544*	148.94187		

* See the note for Table 14.

Table 20

Analysis of regression coefficients for factors for estimating total fat production from cumulative fat production for all records in all areas.

Source	df	Sums of Squares	Mean Squares	F
Level of Fat Prod.	14	1.05059	0.07504	9.40351**
Residual	18,530	147.86122	0.00798	
Total	18,544*	148.91181	0.00798	

* See the note for Table 15.

A study of the above tables indicates that Per Cent Fat, Lactation Number, and Season are highly significant for all areas for both milk and fat factors, with the exception of Season for the analysis of fat for areas 3 and 5. For the latter analysis the effect of Season is still significant. Thus Breed, Age, and Season all affect the shape of the lactation estimation curve. Geographic Area was found to be significant only in the case of the analysis for fat in areas 2 and 4. The marginal means for Geographic Area in this case of milk are 0.87916 and 0.87897, for areas 2 and 4, respectively. As can be seen, these are practically the same and hence the zero sum of squares in Table 11. Prior to combining areas 1 and 2, an analysis was performed using areas 1, 2, and 4. In this analysis, the effect of Geographic Area was found to be highly significant for milk, but not significant for fat. In the analysis for fat in this case, the effect of Per Cent Fat was also found to be non-significant. In spite of this, areas 1 and 2 were combined because there were so few records from these areas. It should be remembered that the analyses for Tables 11, 12, 16, and 17 [as well as 13 and 18] are for the cell means only, thus, the means for the cells in areas 1 and 2 are considered in the analysis to have the same weight as those for area 4. Referring to Table 1, it can be seen that there are approximately 4 times as many records in area 2 [second three states] as in area 1 [first three states] and 5 times as many records in area 4 [New Mexico-Arizona] as in area 1. When the first two areas are combined, the numbers of records in the two groups to be

compared are approximately equal. Because of this disparity in weighting, then, it was decided to use the combined analysis given in Tables 11 and 16. It might be noted, however, that the regression coefficients in area 1 were consistently higher than those for area 2.

There is a highly significant interaction between Per Cent Fat and Season for milk and fat in areas 2 and 4 and for milk in areas 3 and 5, which seems to indicate that different breeds [or cows producing different fat percentages] react differently to different seasonal conditions. This effect also appears in both milk and fat analyses for areas 3 and 5 with respect to the interaction between Breed [per cent fat] and Geographic Area. In this particular case, the significant interaction indicates that cows with different percentages of fat will have a different lactation estimation curve in Colorado than those in Utah or Nevada. For fat, this Area by Breed interaction effect seems actually greater than the effect of Season or Breed [see Section 3.2] on the lactation curve. One might also note that in areas 2 and 4 for milk, there is a significant interaction between Geographic Area and Lactation Number. Also in areas 3 and 5 for milk, there is a significant interaction between Lactation Number and Per Cent Fat. In the analysis for fat in areas 2 and 4, there is a significant three-way interaction between Per Cent Fat, Lactation Number and Season. The effects of these interactions will be looked at more closely in the section dealing with the analysis of the means [see Section 3.3].

An examination of Tables 13 and 18 indicates that neither Days Previous Dry nor Days Open seem to have any significant effect on the shape of the lactation estimation curve for either milk or fat.

Tables 14, 15, 19 and 20 indicate that both Level of Milk Production and Level of Fat Production are highly significant for both milk and fat. The fact that there is an extremely large number of degrees of freedom associated with the within or residual sum of squares could indicate that actually there is a very small difference with a very sensitive test. Since these tests of significance converge to significance as the degrees of freedom become infinite, this could very well be the case here. The means of these analyses will be looked at in greater detail in Section 3.3.

3.2. Results of the analyses of the components of variance and discussion. Although this study is concerned with the fixed effects model and thus, one is not primarily concerned with the estimation of the variance components, it was felt that much valuable information could be obtained by studying the "variances" of the constants involved in this finite population to determine the relative importance of the significant effects, i.e., the proportion of the total variance for which they account. Since a fixed effects model is being assumed, the expected mean squares were calculated, equated to their respective mean squares, and solutions obtained for estimates of the components of variance. The entire analysis is based on the means of the cells, hence, the analyses of the components of variance are

considered as the orthogonal case with one observation per cell.

Below are presented the four four-way analyses in terms of their components of variance.

Table 21

Analysis of components of variance for the analysis of milk coefficients for geographic areas 2 and 4. [See Table 11.]

Source	df	Component of Variance	Coded	%
Geographic Area	1	0.0000000	0	0.000
Per Cent Fat	2	0.0001462	1462**	7.065
Lactation Number	5	0.0007350	7350**	35.516
Season	3	0.0002908	2908**	14.052
GA x PF	2	0.0000000	0	0.000
GA x LN	5	0.0000575	575*	2.778
GA x S	3	0.0000000	0	0.000
PF x LN	10	0.0000000	0	0.000
PF x S	6	0.0001708	1708**	8.253
LN x S	15	0.0000450	450	2.174
GA x PF x LN	10	0.0000525	525	2.537
GA x PF x S	6	0.0000000	0	0.000
GA x LN x S	15	0.0000267	267	1.290
PF x LN x S	30	0.0001150	1150	5.557
Residual	30	0.0004300	4300	20.778
Total	143	0.0020695	20695	100.000

Note that negative components of variance are considered to be zero and are reproduced here as such.

Table 22

Analysis of components of variance for the analysis of milk coefficients for geographic areas 3 and 5. [See Table 12].

Source	df	Component of Variance	Coded	%
Geographic Area	1	0.0000068	68	0.391
Per Cent Fat	4	0.0002835	2835**	16.320
Lactation Number	5	0.0005563	5563**	32.025
Season	3	0.0003868	3868**	22.267
GA x PF	4	0.0000296	296*	1.704
GA x LN	5	0.0000070	70	0.404
GA x S	3	0.0000000	0	0.000
PF x LN	20	0.0000463	463*	2.665
PF x S	12	0.0000508	508**	2.924
LN x S	15	0.0000000	0	0.000
GA x PF x LN	20	0.0000000	0	0.000
GA x PF x S	12	0.0000000	0	0.000
GA x LN x S	15	0.0000500	500	2.878
PF x LN x S	60	0.0000000	0	0.000
Residual	60	0.0003200	3200	18.422
Total	239	0.0017371	17371	100.000

See note on Table 21.

Table 23

Analysis of components of variance for the analysis of fat coefficients for geographic areas 2 and 4. [See Table 16].

Source	df	Component of Variance	Coded	%
Geographic Area	1	0.0000299	299*	1.278
Per Cent Fat	2	0.0000533	533**	2.278
Lactation Number	5	0.0010883	10883**	46.513
Season	3	0.0000481	481**	2.056
GA x PF	2	0.0000395	395	1.688
GA x LN	5	0.0000300	300	1.282
GA x S	3	0.0000128	128	0.547
PF x LN	10	0.0000662	662	2.829
PF x S	6	0.0001342	1342**	5.736
LN x S	15	0.0000100	100	0.427
GA x PF x LN	10	0.0000825	825	3.526
GA x PF x S	6	0.0000317	317	1.355
GA x LN x S	15	0.0000133	133	0.567
PF x LN x S	30	0.0002200	2200*	9.403
Residual	30	0.0004800	4800	20.515
Total	143	0.0023398	23398	100.000

Table 24

Analysis of components of variance for the analysis of fat coefficients for geographic areas 3 and 5. [See Table 17].

Source	df	Component of Variance	Coded	%
Geographic Area	1	0.0000036	36	0.034
Per Cent Fat	4	0.0000350	350**	0.329
Lactation Number	5	0.0099250	99250**	93.388
Season	3	0.0000180	180*	0.169
GA x PF	4	0.0001121	1121**	1.055
GA x LN	5	0.0000000	0	0.000
GA x S	3	0.0000000	0	0.000
PF x LN	20	0.0000000	0	0.000
PF x S	12	0.0000000	0	0.000
LN x S	15	0.0000000	0	0.000
GA x PF x LN	20	0.0000000	0	0.000
GA x PF x S	12	0.0000000	0	0.000
GA x LN x S	15	0.0000240	240	0.226
PF x LN x S	60	0.0000000	0	0.000
Residual	60	0.0005100	5100	4.799
Total	239	0.0106277	106277	100.00

See note on Table 21.

From these tables, it can be seen that the component of variance for Lactation Number accounts for more of the variation than any other source for the entire set of analyses for fat. In areas 3 and 5, it accounts for over 93% of the variation for fat. Lactation Number accounts for a much larger proportion of the variation in fat than it does for milk. There is a larger difference between proportions for milk and fat in areas 3 and 5 than in areas 2 and 4. Per Cent Fat and Season account for a larger portion of the variation for milk than they do for fat. It would seem, therefore, that Per Cent Fat and Season have a greater effect on milk coefficients than they do on fat coefficients. In other words, Per Cent Fat [breed] and Season

exert a larger effect on the lactation estimation curves for milk than they do on those for fat, and Lactation Number exerts a larger effect on the Lactation estimation curve for fat than it does for milk, although the over-all affect of Lactation Number is larger than for any other variables for both milk and fat.

3.3. Results of the analysis of means and discussion. Graphs of the significant variables and interactions are presented below. Tables of comparisons of means using Tukey's[4] hsd test for these effects are given below.

Table 25

Table of means for the significant effects in Table 12, geographic areas 2 and 4 for milk, using Tukey's hsd test.

Per Cent Fat

$s_{\bar{x}} = 0.00299, \quad q_{0.05} = 3.48, \quad \text{hsd} = 0.01045.$

	2	3
	0.88639	0.88612
4 0.86468	0.02171*	0.02144*
	0.00027	

[4]The reason for the use of Tukey's test instead on one of the other available tests for multiple comparisons of means is explained in Section 1.4.3. It was used because of the equal subclasses involved in the analyses. Had an analysis of the means for Level of Production for milk or fat been given, Sheffe's test would have been the ideal test to use, because of the unequal subclass numbers involved. It should be noted that the α-level [Type I error probability] here is 0.05 for each level and not for the over-all effect considered. Thus in Table 26, the probability of an error of Type I is 0.05 for Per Cent Fat for the difference between level 4 and level 2. It is also 0.05 for each of the other differences represented for the levels of Per Cent Fat. If all three comparisons for Per Cent Fat are taken together, the alpha-level is considerably greater.

Table 25

(Continued)

Table of means for the significant effects in Table 12, geographic areas 2 and 4 for milk, using Tukey's hsd test.

Lactation Number

$s_{\bar{x}} = 0.00423, \quad q_{0.05} = 4.30, \quad hsd = 0.01819.$

	1	6	5	2	4
	0.93431*	0.87402	0.87085	0.86268	0.86292
3 0.86261	0.07170*	0.01141	0.00824	0.00707	0.00031
4 0.86292	0.07139*	0.01110	0.00793	0.00676	
2 0.86968	0.06463*	0.00434	0.00117		
5 0.87085	0.06346*	0.00317			
6 0.87402	0.06029*				

Season

$s_{\bar{x}} = 0.00346, \quad q_{0.05} = 3.84, \quad hsd = 0.01329.$

	1	4	2
	0.90067	0.88546	0.86708
3 0.86305	0.03762*	0.02241*	0.00403
2 0.86708	0.03359*	0.01838*	
4 0.88546	0.01521		

Table 26

Table of means for the significant effects in Table 13, geographic areas 3 and 5 for milk, using Tukey's hsd test.

Per Cent Fat

$s_{\bar{x}} = 0.00258, \quad q_{0.05} = 3.98, \quad hsd = 0.01027.$

	2	3	4	5
	0.88371	0.87928	0.86382	0.85300
6 0.84351	0.04020*	0.03577*	0.02031*	0.00949
5 0.84931	0.03071*	0.02628	0.01082*	
4 0.86282	0.01989*	0.01546		
3 0.87928	0.00443			

Table 26

(Continued)

Table of means for the significant effects in Table 13, geographic areas 3 and 5 for milk, using Tukey's hsd test.

Lactation Number

$s_{\bar{x}} = 0.00283$, $q_{0.05} = 4.16$, hsd = 0.01177.

	1 0.91050	6 0.86924	2 0.85859	5 0.85188	3 0.84931
4 0.84846	0.06204*	0.02078*	0.01013	0.00342	0.00085
3 0.84931	0.06119*	0.01993*	0.00928	0.00257	
5 0.85188	0.05862*	0.01736*	0.00671		
2 0.85859	0.05191*	0.01065			
6 0.86924	0.04126				

Season

$s_{\bar{x}} = 0.00231$, $q_{0.05} = 3.74$, hsd = 0.00864.

	1 0.88474	4 0.87699	2 0.85542
3 0.84151	0.04323*	0.03548*	0.01391*
2 0.85542	0.02932*	0.02157*	
4 0.87699	0.00775		

Table 27

Table of means for the significant effects in Table 17, geographic areas 2 and 4 for fat, using Tukey's hsd test.

Geographic Area

$s_{\bar{x}} = 0.00258$, $q_{0.05} = 2.89$, hsd = 0.00746.

$\overset{4}{0.87484} - \overset{2}{0.86629} = 0.00855^*$

Per Cent Fat

$s_{\bar{x}} = 0.00316$, $q_{0.05} = 3.48$, hsd = 0.01100.

	4 0.87598	3 0.87428
2 0.86143	0.01455*	0.01285*
3 0.87428	0.00170	

Table 27

(Continued)

Table of means for the significant effects in Table 17, geographic areas 2 and 4 for fat, using Tukey's hsd test.

Lactation Number

$s_{\bar{x}} = 0.00447$, $q_{0.05} = 4.30$, hsd = 0.01922.

	1	2	4	5	6
	0.93789	0.86578	0.85606	0.85570	0.85541
3 0.85251	0.08538*	0.01327	0.00355	0.00319	0.00290
6 0.85541	0.08248*	0.01037	0.00065	0.00029	
5 0.85570	0.08219*	0.01008	0.00036		
4 0.85606	0.08183*	0.00972			
2 0.86578	0.07211*				

Season

$s_{\bar{x}} = 0.00365$, $q_{0.05} = 3.84$, hsd = 0.01402.

	3	4	1
	0.87847	0.87493	0.86810
2 0.86074	0.01773*	0.01419*	0.00736
1 0.86810	0.01037	0.00683	
4 0.87493	0.00354		

Table 28

Table of means for the significant effects in Table 18, geographic areas 3 and 5 for fat, using Tukey's hsd test.

Per Cent Fat

$s_{\bar{x}} = 0.00326$, $q_{0.05} = 3.98$, hsd = 0.01297.

	5	4	6	3
	0.87182	0.86323	0.86233	0.85694
2 0.85432	0.01750*	0.00891	0.00801	0.00262
3 0.85694	0.01488*	0.00629	0.00539	
6 0.86233	0.00949	0.00090		
4 0.86323	0.00859			

Table 28
(Continued)

Table of means for the significant effects in Table 18, geographic areas 3 and 5 for fat, using Tukey's hsd test.

Lactation Number

$s_{\bar{x}} = 0.00357$, $q_{0.05} = 4.16$, hsd = 0.01485

	1	2	6	5	3
	0.92421	0.86179	0.85635	0.84374	0.84297
4 0.84130	0.08291*	0.02049*	0.01515*	0.00244	0.00167
3 0.84297	0.08124*	0.01882*	0.01338*	0.00077	
5 0.84374	0.08047*	0.01805*	0.01261		
6 0.85635	0.06786*	0.00544			
2 0.86179	0.06242*				

Season

$s_{\bar{x}} = 0.00292$, $q_{0.05} = 3.74$, hsd = 0.01092.

	3	4	1
	0.86691	0.86531	0.85811
2 0.85658	0.01033	0.00873	0.00153
1 0.85811	0.00880	0.00720	
4 0.86531	0.00160		

Beginning with Per Cent Fat for milk in areas 2 and 4, [Table 25], it can be seen that levels 2 and 3 differ significantly from level 4, but not from each other. This would suggest that the lactation estimation curves for milk in areas 2 and 4 for low fat producing Holsteins and Holsteins and Milking Shorthorns are essentially the same, but that they differ from the average curve for the other breeds. This relationship is given, diagramatically, in Figure 2 below.

Figure 2

A diagram of the mean comparisons for milk for Per Cent Fat in areas 2 and 4.

23   4

This would suggest using only two levels of this factor, i.e., one for high producing cows [4.1% and over] and one for low producing cows [4.0% and under].

The analysis of means for Lactation Number in areas 2 and 4 for milk substantiate the results found by Madden <u>et al</u>. [1956] and other workers cited in Section 1.2.4, namely, that factors for milk can be separated into two groups, i.e., first lactations and second lactations or greater. This point is illustrated in the following diagram.

<u>Figure 3</u>

A diagram of the mean comparisons for milk for Lactation Number in areas 2 and 4.

1 <u>6 5 2 4 3</u>

The analysis of means for Season indicates that curves for cows freshening in the spring and summer can be combined, giving three groups, one for winter, one for spring and summer, and one for fall. Diagramatically, this relationship is given in Figure 4.

<u>Figure 4</u>

A diagram of the mean comparisons for milk for Season in areas 2 and 4.

1 4 <u>2 3</u>

Cows freshening in winter have a steeper lactation estimation curve than those freshening during other seasons of the year, and cows freshening in cold seasons have a steeper lactation estimation curve than do those freshening in hot months. [See

figure 14 and the discussion following].

In the analysis of means for areas 3 and 5 for milk [Table 26], for Per Cent Fat, levels 2 and 3 and levels 5 and 6 could be combined, but the other levels differ significantly from each other. This can be seen in the following diagram.

Figure 5

A diagram of the mean comparisons for milk for Per Cent Fat in areas 3 and 5.

$\underline{2 \quad 3} \qquad\qquad 4 \quad \underline{5 \qquad\qquad 6}$

In areas 3 and 5, the lactation estimation curves for milk could be considered different for breeds corresponding to 4.0% and under, 4.1% to 4.6%, and 4.7% and over, respectively. In areas 2 and 4 cows producing 4.7% to 5.2% combined with the higher fat producing cows and in areas 3 and 5 they are not. In the former case, results similar to those for areas 3 and 5 would probably have been obtained also, had the 4.7% to 5.2% group been kept separate.

For Lactation Number, a more complicated relationship is found [see Figure 13]. Here, first lactations differ from all other lactations, but curves for sixth lactations or over also differ significantly from those for all other lactations [with the exception of the second lactations].

Figure 6

A diagram of the mean comparisons for milk for Lactation Number in areas 3 and 5.

$1 \qquad\qquad\qquad \underline{6 \qquad\qquad 2} \qquad 5 \quad 3\,4$

In this case, the curves could probably be grouped into first lactations, adult cow lactations, and old cow lactations. Or they could be grouped so that the second and sixth lactations are estimated by the same factor.

A different relationship is found for different seasons in areas 3 and 5 for milk. The analysis of means indicates that curves for cows freshening in the fall and winter could be combined, but that they differ for cows freshening in the spring and summer. See the diagram below.

Figure 7

A diagram of the mean comparisons for milk for Season in areas 3 and 5.

1 4 2 3

It can be seen that cows freshening in the cold seasons have a steeper lactation estimation curve than those freshening in the hot seasons, with the steepest curve being for cows freshening in the winter months. [See Figure 14].

Cows from New Mexico and Arizona have steeper curves than do those from the northern states [Table 27].

For Per Cent Fat, for fat an inverse relationship to that noted for milk in areas 2 and 4 is found. The Holsteins and the red breeds should be grouped together and the low fat producers kept separate. Again a diagram of the mean comparisons is given.

Figure 8

A diagram of the mean comparisons for fat for Per Cent Fat in areas 2 and 4.

4 3 2

The analysis of means for Lactation Number gives the same result, in terms of significance, for fat as for milk, although the rankings of the means are different. Here again, one need only have a factor for the first lactations and one for all succeeding lactations.

A comparison of the means for Season in areas 2 and 4 for fat, indicates a somewhat more complicated situation than that for milk; the following diagram may serve to illustrate the situation.

Figure 9

A diagram of the mean comparisons for fat for Season in areas 2 and 4.

3 4 1 2

This diagram of the analysis of means indicates that lactation estimation curves for summer, fall, and winter do not differ from each other, and that the curves for winter and spring do not differ from each other. It would probably be wisest to combine summer and fall, and to combine winter and spring. It should be noted that the steepest curve is for cows freshening in the summer, and the flattest curve is for cows freshening in the spring [in terms of estimating fat production].

Table 28, which gives the comparisons of means for fat for areas 3 and 5 yields the following diagram for Per Cent Fat.

Figure 10

A diagram of the mean comparisons for fat for Per Cent Fat in areas 3 and 5.

5 4 6 3 2

This diagram of the analysis of means indicates that while the average curve for 4.7% to 5.2% fat producers differs from those for 3.4% and under and 3.5% to 4.0% fat producers, the curve for 4.7% to 5.2% does not differ from those of 5.3% and over and 4.1% to 4.6% fat producers, nor do these latter two differ from those of 3.4% and under and 3.5% to 4.0% fat producers. This abrupt departure from the linear trend which existed in the milk means as the Per Cent Fat increased, strengthens the idea that breed differences rather than average per cent fat affect the shape of the lactation estimation curve. It seems more likely that if the effect were due to level of per cent fat, the trend would be more linear [or slightly curvelinear in nature], and not have the tremendous drop in the sixth level, which was encountered.

The relationship of the means with respect to Lactation Number seems somewhat different in the fat coefficients than in the milk coefficients [Figure 11].

<u>Figure 11</u>

A diagram of the mean comparisons for fat for Lactation Number in areas 3 and 5.

1 <u>2 6</u> <u>5 3 4</u>

This diagram brings out the fact that if the means are combined as for milk, the groupings would be the same, with one factor being used to estimate the first lactation production, one for the second and sixth or greater lactations, and one for the intermediate lactations. The only differences between

these two tables [Table 26 and Table 28] are that the rankings of the means for the second and sixth lactations are reversed, and the difference between the second and fourth lactation means is not significant for milk as it is for fat. Thus, the same pattern of estimation for milk and fat could be used in structuring the factors.

The analysis of means for Season did not yield any significant differences, however, here as in areas 2 and 4, curves for summer and fall seem steeper than for winter and spring.

Graphs of the various levels will now be given as a further aid in the study of the means of significant effects.

FIGURE 12

GRAPH OF AVERAGE PER CENT FAT FOR MILK AND FAT FOR BOTH SETS OF AREAS.
FACTORS ARE CODED TO .91000 = 900, ETC.

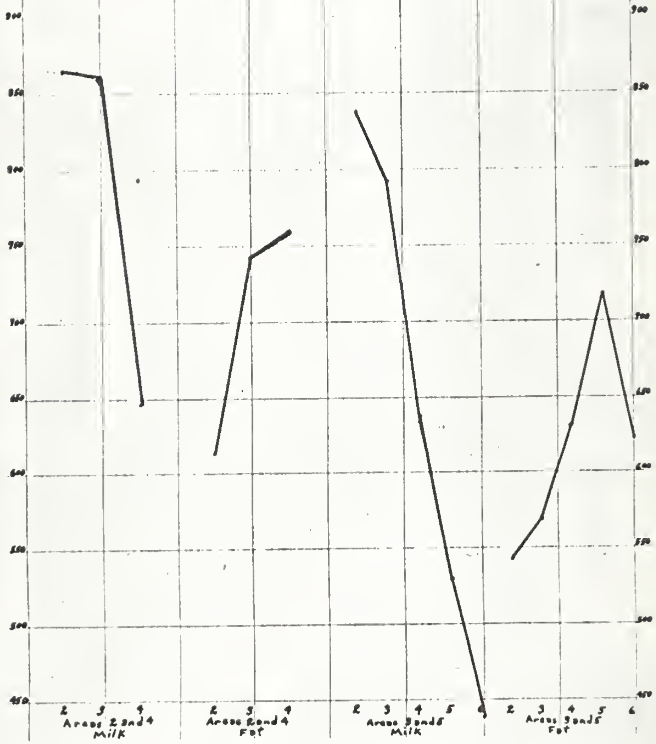


FIGURE 19

GRAPH OF LACTATION NUMBER FOR MILK AND FAT FOR BOTH SETS OF AREAS.

FACTORS ARE CODED TO .81000 = 800, ETC.

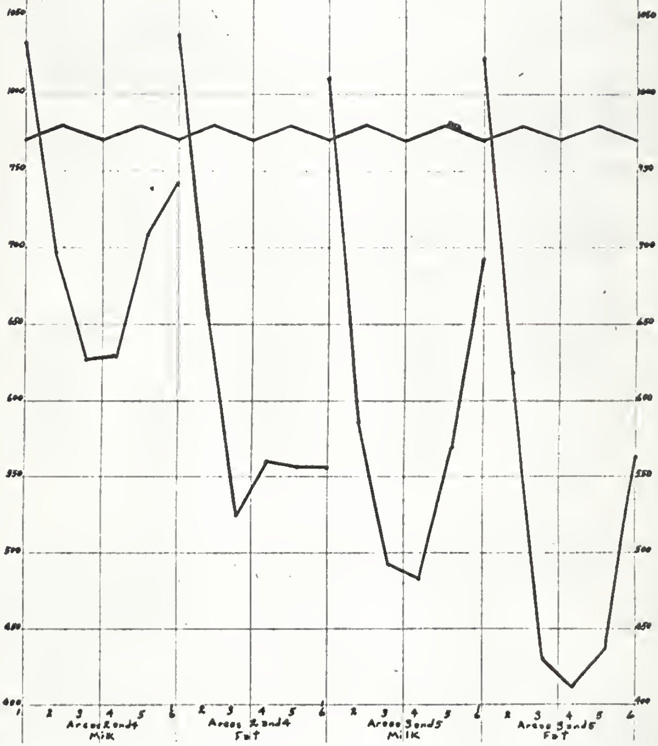


FIGURE 14

GRAPH OF SEASON FOR MILK AND FAT FOR BOTH SETS OF AREAS
FACTORS ARE CODED TO .95000 = 950, ETC.

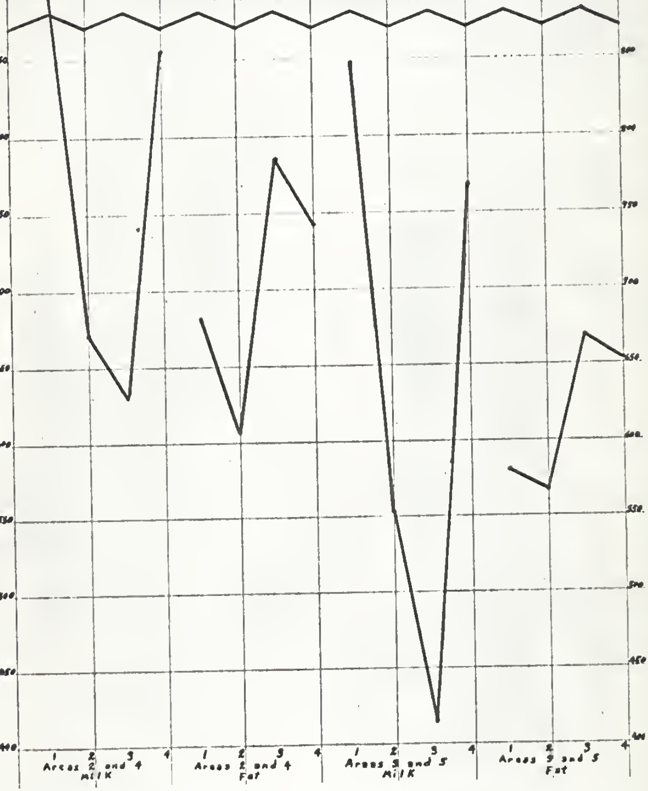
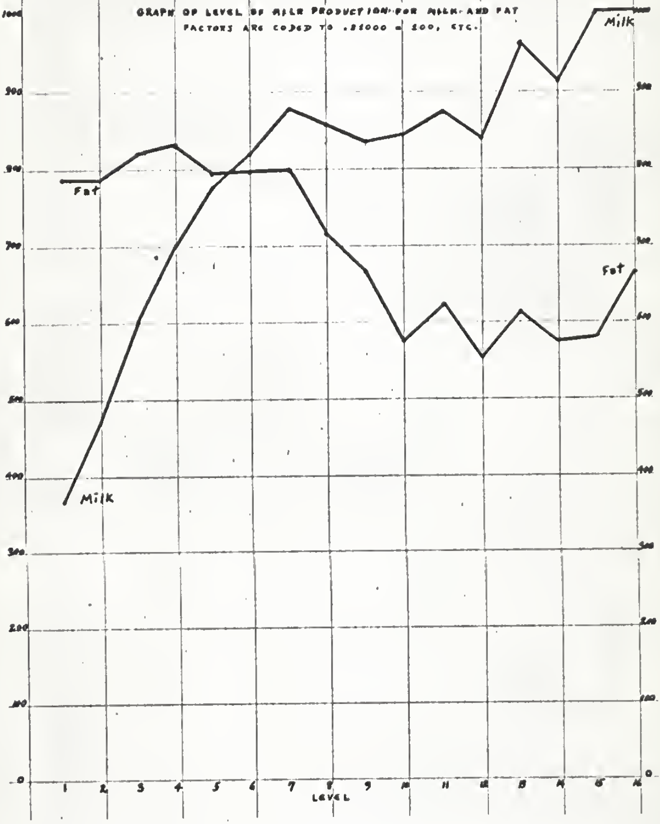


FIGURE 15

GRAPH OF LEVEL OF MILK PRODUCTION FOR MILK AND FAT
FACTORS ARE CODED TO .01000 = 100, ETC.



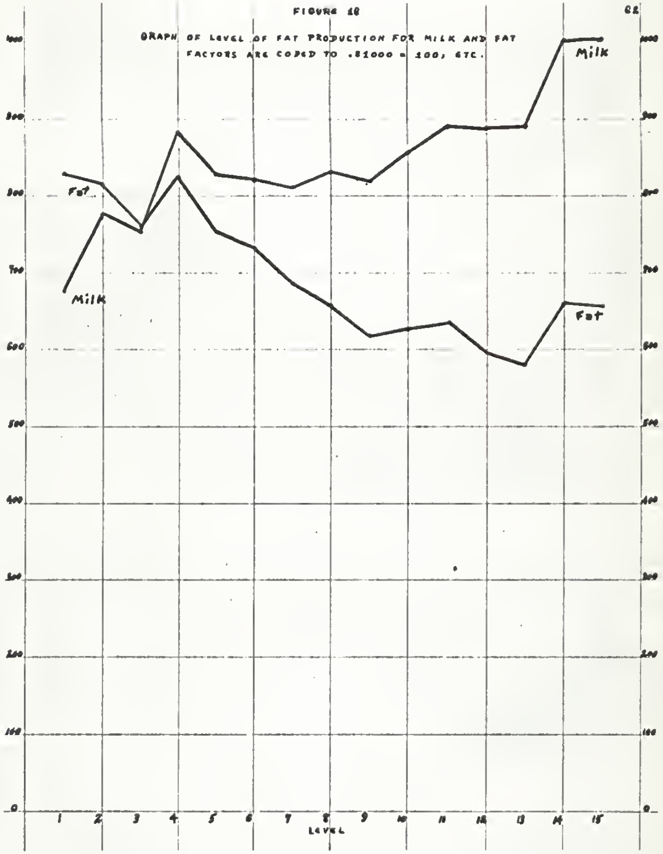


FIGURE 18

GRAPH OF LEVEL OF FAT PRODUCTION FOR MILK AND FAT FACTORS ARE CODED TO .81000 = 100, ETC.

The graphs in Figures 12 through 16 show the effects of moving from one level of each factor to another in terms of the average or marginal factors. On each of these graphs, a high coefficient indicates a steep lactation estimation curve which means that the lactation curve for the cow is quite flat. This one coefficient characterizes the entire lactation estimation curve.

It can be seen from Figure 12 that the coefficients for Average Per Cent Fat for milk in both areas 2 and 4 and areas 3 and 5 show a steep decline as one moves from low fat producing cows to high producing cows. This would indicate that low fat producing breeds produce proportionately more milk later in the lactation than they do in the early part; the opposite is true for the high fat producing breeds. As would be expected, coefficients for these two sets of areas for fat indicate that as Average Per Cent Fat increases, the coefficients also increase. This indicates that lactation estimation curves are steeper in high fat producing breeds than in low fat producing breeds. This means that proportionately more fat is produced early in the lactation in low fat producing breeds than in high fat producing breeds. Both Smith [1959] and Johansson [1961] have pointed out that as per cent fat increases, level of production decreases. Note that the magnitude of this effect is greater for milk coefficients than for fat coefficients. This indicates that when per cent fat is high, milk production is low, thus masking the effects of fat production. The sharp drop between the fifth and sixth levels [5.3% to 5.8% and 5.9% and over] for

fat in areas 3 and 5 could possible be indicative of a breed difference, rather than a difference due to change in level of per cent fat, as has been mentioned earlier.

From Figure 13, it can be seen that the effect of lactation number is similar whether for milk or fat, or whether in areas 2 and 4 or 3 and 5. From the graphs, there seems to be a relatively larger effect in areas 3 and 5 than in areas 2 and 4. This could be due to actual differences in the effect for the different areas, or it could be due to the fact that more records were used in the calculation of means in areas 3 and 5 than in areas 2 and 4. Another phenomenon to be noted is that for milk coefficients, the coefficient for the sixth and greater lactations is consistently greater than that for second lactations. For fat coefficients, the opposite is true. Each of these curves seems to illustrate the fact that the shape of the lactation curve changes as a cow matures in much the same manner as the total production changes. The greatest change occurs between the first and succeeding lactations, the steepest lactation estimation curve being for first lactations. The flattest lactation estimation curves are generally for third, fourth, and fifth lactations. Except for the case of fat for areas 2 and 4, the effect seems to be distinctly quadratic in nature, meaning that older cows have curves which are becoming more and more like those of first lactation cows. The effect, as expressed by the graph of these coefficients is very similar to the effect of age on total production. These graphs indicate that for first lactation cows, a larger

proportion of milk and fat is produced later in the lactation than in the beginning portion. This proportion decreases at a decreasing rate until the third and fourth lactations, after which it again begins to increase.

Although the graphs shown in Figure 14 must be interpreted in terms of individual points for winter, spring, summer, and fall since their consecutiveness is not the important fact, it can be seen that the graphs for milk are similar in both sets of areas, as are those for fat. The coefficients indicate that curves for fall and winter are similar and the curves for spring and summer are similar for milk, and that curves for winter and spring are similar and those for summer and fall are similar for fat. This has not been mentioned in earlier studies. In areas 2 and 4, the coefficients seem to be higher, in general, for both milk and fat than those in areas 3 and 5. The coefficient for milk is highest in the winter and lowest in the summer, and the coefficient for fat is highest in the summer and lowest in the spring.

In Figure 15, it can be seen that the coefficients for milk follow a somewhat curvelinear trend which increases as one moves from low to high producing cows. An opposite effect is noted for fat, although the effect is not so pronounced. If a smooth curve is fitted through the points on the graph, it appears that both effects are cubic in nature. The coefficients for level of fat production in Figure 16 exhibit the same shape curves for milk and fat as those in Figure 15; the same sort of cubic effect is represented. The irratic

scatter of points is probably due to random error.

3.4. Results of the analyses of interactions and discussion. Below are presented tables of significant interactions for the four four-way analyses. Following these are graphs representing each table.

Table 29

Table of interactions for Geographic Area by Lactation Number for milk for areas 2 and 4. [See Table 11.]

		LN						
	LEVEL	1	2	3	4	4	6	AVE
GA	2	0.92290	0.87363	0.86093	0.87172	0.87319	0.87256	0.87916
	4	0.94573	0.86573	0.86428	0.85412	0.86850	0.87548	0.87897
	AVE	0.93431	0.86968	0.86261	0.86292	0.87085	0.87402	0.87906

Table 30

Table of interactions for Per Cent Fat by Season for milk for areas 2 and 4. [See Table 11.]

		S				
	LEVEL	1	2	3	4	AVE
PF	2	0.89678	0.87199	0.88358	0.89322	0.88639
	3	0.89939	0.87251	0.87670	0.89589	0.88612
	4	0.90586	0.85673	0.82888	0.86272	0.86468
	AVE	0.90067	0.86708	0.86305	0.88546	0.87906

Table 31

Table of interactions for Geographic Area by Per Cent Fat for milk for areas 3 and 5. [See Table 12.]

		PF					
	LEVEL	2	3	4	5	6	AVE
GA	3	0.88824	0.88375	0.86734	0.85739	0.83742	0.86683
	5	0.87918	0.87480	0.86030	0.84860	0.84960	0.86250
	AVE	0.88371	0.87928	0.86383	0.85300	0.84351	0.86466

Table 32

Table of interactions for Per Cent Fat by Lactation Number for milk for areas 3 and 5. [See Table 12].

		LN						
	LEVEL	1	2	3	4	5	6	AVE
	2	0.94251	0.87417	0.86616	0.86626	0.86989	0.88328	0.88371
	3	0.93164	0.87312	0.86143	0.86271	0.87011	0.87666	0.87928
PF	4	0.90926	0.86456	0.85285	0.84800	0.84406	0.86420	0.86382
	5	0.89621	0.84142	0.83506	0.85029	0.84436	0.85064	0.85300
	6	0.87288	0.83965	0.83104	0.81506	0.83098	0.87143	0.84351
	AVE	0.91050	0.85859	0.84931	0.84846	0.85188	0.86924	0.86466

Table 33

Table of interactions for Per Cent Fat by Season for milk for areas 3 and 5. [See Table 12].

		S				
	LEVEL	1	2	3	4	AVE
	2	0.88887	0.87524	0.87476	0.89598	0.88371
	3	0.89578	0.87418	0.86076	0.88641	0.87928
PF	4	0.88805	0.84752	0.83640	0.88332	0.87928
	5	0.88065	0.84477	0.82669	0.85988	0.85300
	6	0.87035	0.83540	0.80893	0.85935	0.84351
	AVE	0.88474	0.85542	0.84151	0.87699	0.86466

Table 34

Table of interactions for Per Cent Fat by Season for fat for areas 2 and 4. [See Table 16].

	S				
LEVEL	1	2	3	4	AVE
2	0.85370	0.83871	0.88133	0.87197	0.86143
3	0.86562	0.86463	0.88570	0.88115	0.87428
4	0.88500	0.87889	0.86837	0.87166	0.87598
AVE	0.86810	0.87074	0.87847	0.87493	0.87056

Table 35

Table of interactions for Geographic Area by Per Cent Fat for fat for areas 3 and 5. [See Table 17].

		PF					
LEVEL		2	3	4	5	6	AVE
	3	0.85713	0.86424	0.87297	0.87319	0.85098	0.86370
GA	5	0.85152	0.84965	0.85350	0.87044	0.87367	0.85976
AVE		0.85432	0.85694	0.86323	0.87182	0.86233	0.86173

FIGURE 17
[SEE TABLE 29]

GRAPH OF INTERACTIONS FOR GEOGRAPHIC AREA X LACTATION NUMBER FOR MILK FOR AREAS 2 AND 4.

FACTORS ARE CODED TO .81000 = 400, ETC.

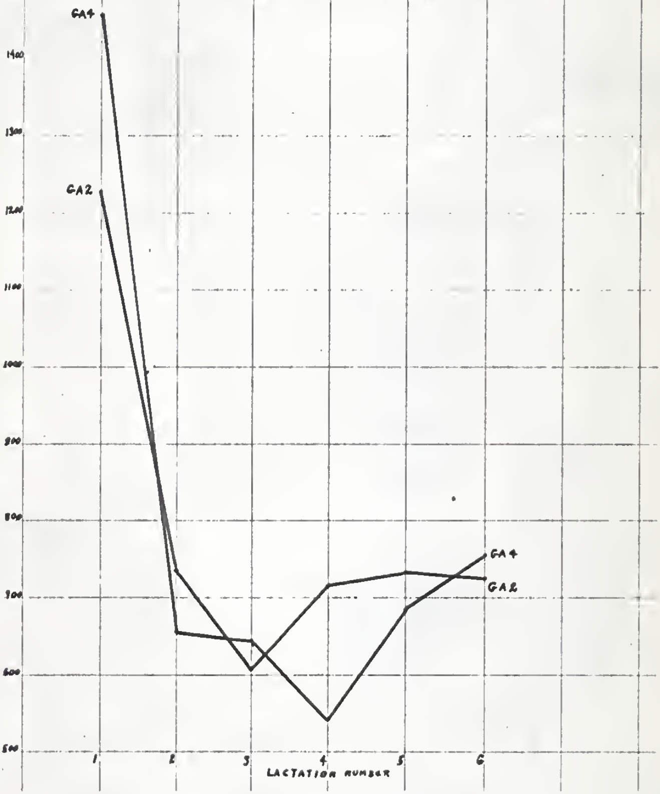


FIGURE 18
[SEE TABLE 30]

GRAPH OF INTERACTIONS FOR PER CENT FAT x SEASON FOR MILK FOR AREAS 1 AND 4.
FACTORS ARE CODED TO .81000 = 100, ETC.

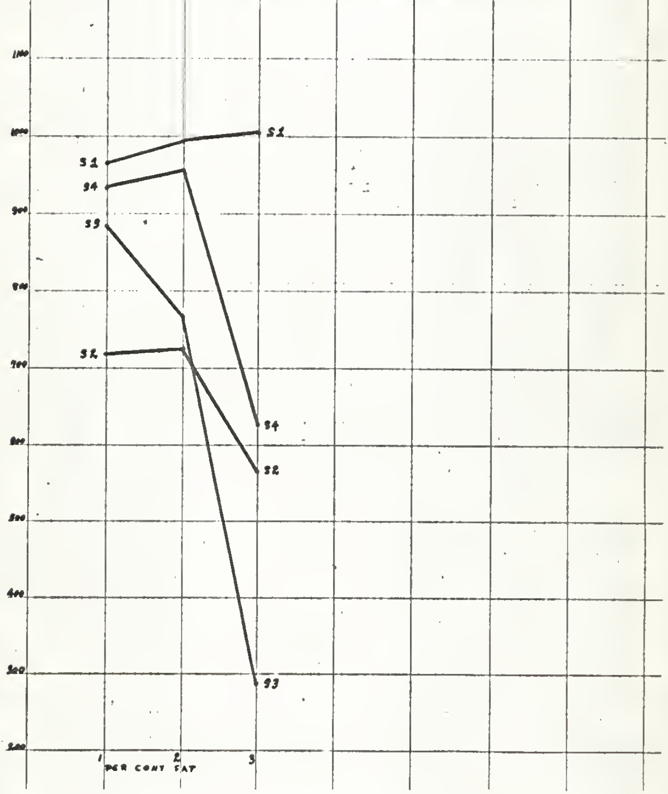


FIGURE 19

[SEE TABLE 31]

GRAPH OF INTERACTIONS FOR GEOGRAPHIC AREA x PER CENT FAT FOR MILK FOR AREAS 3 AND 5

FACTORS ARE CODED ON .$1000 = 100, ETC.

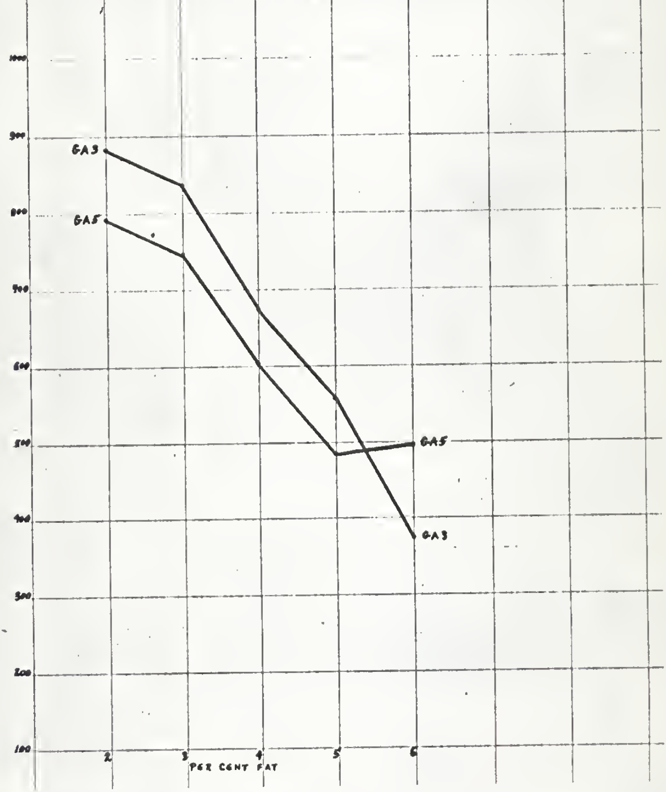


FIGURE 20
[SEE TABLE 32]

GRAPH OF INTERACTIONS FOR PER CENT FAT × LACTATION NUMBER FOR MILK FOR AREAS 3 AND 5.
FACTORS ARE CODED ON .01000 = 600, ETC.

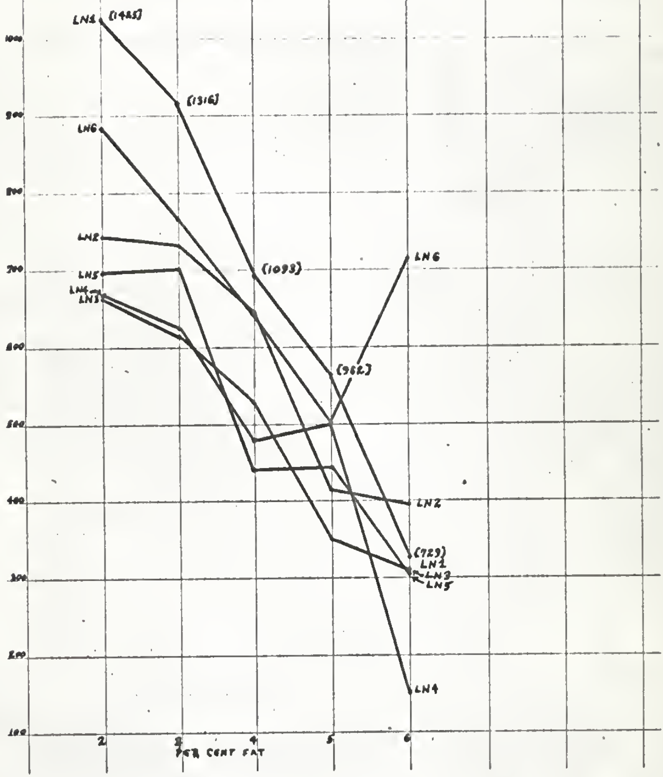


FIGURE 21
[SEE TABLE 32]

GRAPH OF INTERACTIONS FOR PER CENT FAT x SEASON FOR MILK FOR AREAS 3 AND 5.
FACTORS ARE CODED ON .$1000 = 100; ETC.

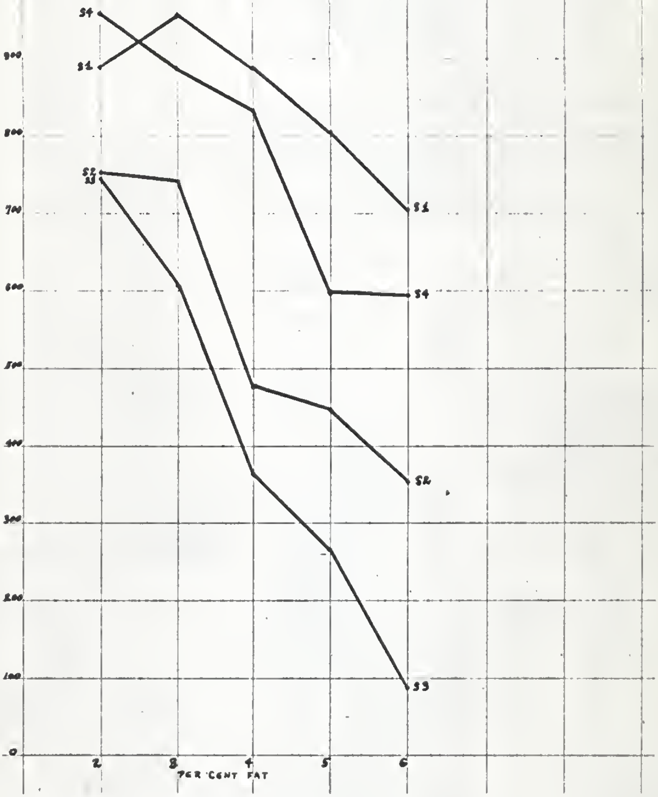


FIGURE 22
[SEE TABLE 34]

GRAPH OF INTERACTIONS FOR PER CENT FAT & SEASON FOR FAT FOR AREAS 2 AND 4.
FACTORS ARE CODED ON .21000 = 200, ETC.

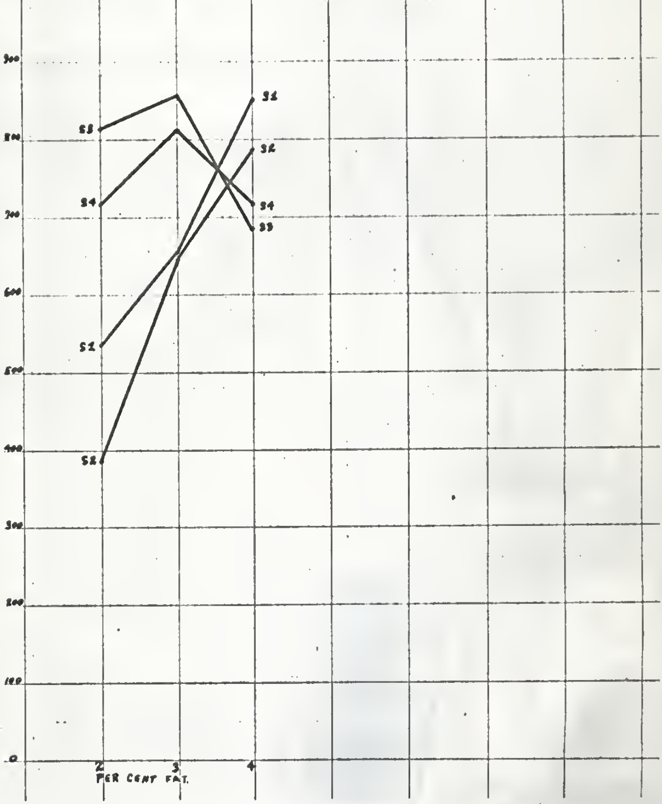
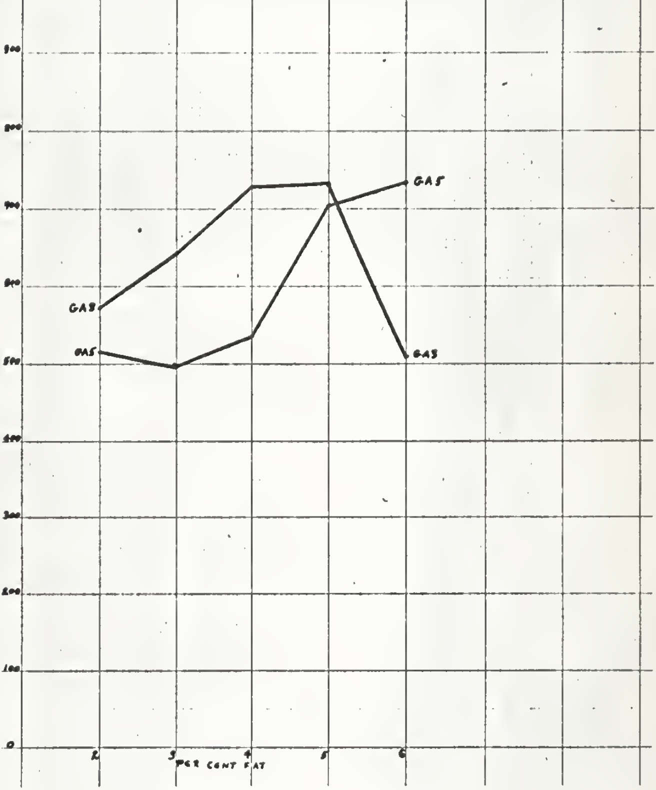


FIGURE 23
[SEE TABLE 35]

GRAPH OF INTERACTIONS FOR GEOGRAPHIC AREA × PER CENT FAT FOR FAT FOR AREAS 3 AND 5
FACTORS ARE CODED TO .81000 = 860, ETC.



The above graphs and tables point out the nature of the interactions. A study of Figure 17 indicates that certain differences do exist between the behavior of Lactation Number in Area 2 and that in Area 4. The difference between first lactations in the two areas indicates that New Mexico-Arizona cows have a steeper lactation estimation curve than those in Area 2, but in terms of interaction, it is important to note that the ranks of the coefficients for the two areas reverse at every level of Lactation Number except one. The steepness of the lactation estimation curve is a minimum at the third lactation for Area 2 and at the fourth lactation for Area 4. Also, the over-all effect of Lactation Number is greater in Area 4 than in Area 2.

From Figure 18, one can see a very definite difference between the effects of level of Average Per Cent Fat, depending on the season in which the cow freshens. For those that freshen in winter [December-February], the higher the per cent fat, the higher the proportion of milk the cow produces later in her lactation, although the effect is not as pronounced as those for the other three seasons. This would indicate that a cow which produces a high percentage of fat in her milk, freshening in winter, would produce the largest proportion of her milk in the summer and fall months. A high level of Per Cent Fat in a cow freshening in the summer [June-August] indicates that she produces a higher proportion of her milk during these months. Coefficients for high fat cows are consistently lower than for low fat cows. These abrupt differences seem especially

indicative of breed differences, rather than differences due to per cent fat. In Figure 19, this difference in high producing cows again appears [though this time not in terms of season]. Colorado [Area 3] has a higher proportion of colored breeds than does Utah-Nevada [Area 5], hence, one could say that the graph for GA5 probably indicates effects of per cent fat change, and the significant drop between PF5 [4.7-5.2%] and PF6 [5.3% and over] would indicate a breed difference. The other points are roughly parallel.

The graph in Figure 20 seems to indicate that the depressing effect on steepness of the lactation estimation curve caused by an increase in level of Per Cent Fat has much more effect on young cows than on old ones. The abrupt change in trend displayed by PF5 and PF6 could also indicate a breed difference, as could the lack of smoothness also noted in the graphs between PF3, PF4, and PF5.

The abrupt drop noted in Figure 18 does not show up in Figure 21. This could be because of the fact that since Utah-Nevada is predominantly Holstein what red-breed effect does come from Colorado is masked by the Utah-Nevada Holsteins. It is interesting to note the change in S1 between PF2 and PF3.

From Figure 22, a Season effect can be seen interacting with Per Cent Fat. As has been earlier pointed out, the effect of Per Cent Fat on the milk lactation estimation curve is practically the opposite of the effect on the fat lactation estimation curve. Instead of increase of level of Per Cent Fat having a depressing effect on the steepness of the

lactation estimation curve, it actually has the opposite effect. Mean separation procedures here indicated that Seasons 1 and 2 should be combined, as should Seasons 3 and 4. In this graph, it can be seen that these separations still are indicated. All the interaction effect is with respect to the high fat cows. The fact that Season interacts so strongly in this case is again indicative of the breed effect.

The effects of area by Per Cent Fat interaction in Figure 23 are similar to those shown in Figure 19. This again is probably indicative of the difference in proportions of the red breeds in Colorado and in Utah-Nevada. There could be a larger distribution of the red breeds in other states in areas 2 and 4, than in areas 3 and 5, and this could probably account for the differences between graphs for interactions with respect to Per Cent Fat for the two sets of areas, and why Per Cent Fat interacted with area for both milk and fat only in areas 3 and 5.

3.5 Factors to be used for the lactation estimation curve.

3.5.1. Preliminaries. The discussion of the first four sections of this chapter indicate that a different structure of factors should be used depending on the area in which the estimates are to be used, and whether they are for milk or fat. The analyses have brought out the fact that there should definitely be a different set of factors for milk than for fat. The discussion of interactions, all of which were in terms of Per Cent Fat, indicate that the fact that Breed is not known tends to confuse the effect of Area, and Area tends to confuse the true interaction of Breed with Season. It appears from the analysis of

Geographic Area by Per Cent Fat, that there is probably an effect due to Per Cent Fat and one due to Breed. Because of the lack of knowledge about Breed and its effect on Per Cent Fat, and the fact that Per Cent Fat and Season interact in most cases, and that Per Cent Fat interacts with some of the other variables, it was decided to leave the factors in their original form for use in estimation until further work could be done with respect to Breed and Per Cent Fat. In using cell means, interaction effects are not confused in any way, and the estimates should not be affected by the significant main or interaction effects. The number of factors is, of course, somewhat unweildy. Factors will now be given in terms of each of the four-four-way analyses, and also for the four analyses in terms of level of production.

3.5.2. Factors for estimating total from part milk and fat production areas 2 and 4. Tables 25 and 27 and Figures 2-4, 8-9, and 12-14 indicate that certain levels of the variables could possibly be combined. The differences in combination indicated that a different structure existed for the variables for milk than for fat. Because of the significant interactions involved, the factors are not combined, but are presented as cell means for all effects.

Table 36

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 2

[North Dakota, Iowa, Nebraska, Montana, Idaho, and Wyoming]

	Average Per Cent Fat-3.4% and under	
	Milk	Fat
First Lactations		
December-February	0.95141	0.90529
March-May	0.93451	0.90576
June-August	0.93231	0.94179
September-November	0.94271	0.92937
Second Lactations		
December-February	0.87682	0.83088
March-May	0.88493	0.85752
June-August	0.87431	0.90300
September-November	0.86054	0.83723
Third Lactations		
December-February	0.85978	0.81853
March-May	0.83907	0.87518
June-August	0.87300	0.87518
September-November	0.87396	0.84391
Fourth Lactations		
December-February	0.94249	0.88141
March-May	0.87040	0.81032
June-August	0.86692	0.89288
September-November	0.90277	0.88121
Fifth Lactations		
December-February	0.85926	0.81845
March-May	0.88440	0.85394
June-August	0.92110	0.87180
September-November	0.87817	0.85891
Sixth Lactations and over		
December-February	0.89862	0.85227
March-May	0.82306	0.78282
June-August	0.87520	0.86582
September-November	0.90789	0.86650
	Average Per Cent Fat-3.5% to 4.0%	
First Lactations		
December-February	0.94585	0.90855
March-May	0.89650	0.88448
June-August	0.90769	0.93700
September-November	0.94451	0.93761

Table 36
[continued]

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 2
[North Dakota, Iowa, Nebraska, Montana, Idaho, and Wyoming]

	Average Per Cent Fat-3.5% to 4.0%	
	Milk	Fat
Second Lactations		
December-February	0.90729	0.88052
March-May	0.88557	0.86101
June-August	0.87604	0.90006
September-November	0.88936	0.87334
Third Lactations		
December-February	0.86328	0.82374
March-May	0.85988	0.80934
June-August	0.84846	0.86863
September-November	0.89517	0.88255
Fourth Lactations		
December-February	0.87483	0.82242
March-May	0.86661	0.82709
June-August	0.83494	0.82746
September-November	0.88084	0.85975
Fifth Lactations		
December-February	0.90930	0.84557
March-May	0.83704	0.81281
June-August	0.86874	0.87902
September-November	0.96311	0.83393
Sixth Lactations and over		
December-February	0.90607	0.82480
March-May	0.86636	0.86008
June-August	0.90036	0.90770
September-November	0.90241	0.87596
	Average Per Cent Fat-4.1% and over	
First Lactations		
December-February	0.94890	0.94258
March-May	0.89103	0.93601
June-August	0.88952	0.93521
September-November	0.88982	0.91253
Second Lactations		
December-February	0.88650	0.84565
March-May	0.85278	0.83405
June-August	0.82361	0.88423
September-November	0.86586	0.86406

Table 36
[continued]

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 2
[North Dakota, Iowa, Nebraska, Montana, Idaho, and Wyoming]

	Average Per Cent Fat-4.1% and over	
	Milk	Fat
Third Lactations		
December-February	0.87980	0.83808
March-May	0.86639	0.87266
June-August	0.84083	0.90124
September-November	0.82609	0.83080
Fourth Lactations		
December-February	0.92603	0.87225
March-May	0.82359	0.88061
June-August	0.83214	0.86496
September-November	0.83908	0.84230
Fifth Lactations		
December-February	0.88592	0.85863
March-May	0.87697	0.92573
June-August	0.80288	0.02606
September-November	0.89142	0.90557
Sixth Lactations and over		
December-February	0.96896	0.90188
March-May	0.80575	0.82959
June-August	0.73842	0.74906
September-November	0.87766	0.87135

Area 4
[New Mexico and Arizona]

	Average Per Cent Fat-3.4% and under	
First Lactations		
December-February	0.97277	0.94521
March-May	0.93617	0.92246
June-August	0.95232	0.97166
September-November	0.97051	0.96907
Second Lactations		
December-February	0.87471	0.83975
March-May	0.86168	0.83855
June-August	0.86758	0.87342
September-November	0.89332	0.88776
Third Lactations		
December-February	0.88339	0.83206
March-May	0.85138	0.83793
June-August	0.84021	0.83022
September-November	0.86386	0.83616

Table 36
[continued]

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 4

[New Mexico and Arizona]

	Average Per Cent Fat-3.4% and under	
	Milk	Fat
Fourth Lactations		
December-February	0.87092	0.83025
March-May	0.84171	0.79888
June-August	0.81703	0.79820
September-November	0.84191	0.84985
Fifth Lactations		
December-February	0.87974	0.83815
March-May	0.89097	0.82997
June-August	0.86508	0.83834
September-November	0.87646	0.83720
Sixth Lactations and over		
December-February	0.89140	0.85214
March-May	0.84556	0.81064
June-August	0.91791	0.91371
September-November	0.90110	0.86652
	Average Per Cent Fat-3.5% to 4.0%	
First Lactations		
December-February	0.97050	0.96472
March-May	0.94917	0.94915
June-August	0.95099	0.96786
September-November	0.96816	0.88362
Second Lactations		
December-February	0.88920	0.88598
March-May	0.86380	0.87910
June-August	0.85449	0.86276
September-November	0.88341	0.88362
Third Lactations		
December-February	0.86330	0.85434
March-May	0.84865	0.84326
June-August	0.87421	0.87608
September-November	0.88856	0.87372
Fourth Lactations		
December-February	0.91909	0.89125
March-May	0.86901	0.89048
June-August	0.86802	0.87390
September-November	0.85949	0.84202

Table 36
[continued]

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 4

[New Mexico and Arizona]

	Average Per Cent Fat-3.5% to 4.0%	
Fifth Lactations	Milk	Fat
December-February	0.87408	0.84922
March-May	0.85918	0.86283
June-August	0.86391	0.85504
September-November	0.88228	0.84904
Sixth Lactations and over		
December-February	0.86984	0.83272
March-May	0.86839	0.89593
June-August	0.87251	0.87294
September-November	0.98333	0.88265
	Average Per Cent Fat-4.1% and over	
First Lactations		
December-February	0.93383	0.94384
March-May	0.90562	0.92525
June-August	0.89666	0.93210
September-November	0.94210	0.96234
Second Lactations		
December-February	0.88480	0.89515
March-May	0.84691	0.85428
June-August	0.83652	0.86191
September-November	0.83235	0.84498
Third Lactations		
December-February	0.89360	0.90485
March-May	0.85200	0.85920
June-August	0.85356	0.87055
September-November	0.85870	0.85799
Fourth Lactations		
December-February	0.86531	0.86761
March-May	0.82002	0.87342
June-August	0.81368	0.90299
September-November	0.86325	0.86401
Fifth Lactations		
December-February	0.89445	0.87111
March-May	0.90703	0.92643
June-August	0.76627	0.82402
September-November	0.86257	0.86510

Table 36
[continued]

Factors for use in estimating total from part milk and fat production in areas 2 and 4.

Area 4

[New Mexico and Arizona]

	Average Per Cent Fat-4.1% and over	
Sixth Lactations and over	Milk	Fat
December-February	0.90226	0.87833
March-May	0.83264	0.82946
June-August	0.85242	0.86811
September-November	0.85836	0.83891

3.5.3. Factors for estimating total from part milk and fat production in areas 3 and 5. Tables 26 and 28 and Figures 5-7, 10-11, and 12-14 indicate that certain levels of the variables could be combined. The differences in combinations again indicate basic differences in structure between milk and fat. The differences between areas 3 and 5 is probably due largely to the difference in relative proportion of breeds between the two areas [see Section 3.4] and to interaction of Breed and Per Cent Fat with other variables, particularly Season. Because of the significant interactions involved, the factors are not combined, but are presented as cell means for all effects.

Table 37

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 3

[Colorado]

	Average Per Cent Fat-3.4% and under	
	Milk	Fat
First Lactations		
December-February	0.94801	0.90233
March-May	0.93690	0.91171
June-August	0.94314	0.93929
September-November	0.97687	0.94981
Second Lactations		
December-February	0.88845	0.85541
March-May	0.87234	0.84906
June-August	0.86611	0.87218
September-November	0.89103	0.85482
Third Lactations		
December-February	0.87545	0.83142
March-May	0.87797	0.84072
June-August	0.86045	0.84365
September-November	0.88624	0.86172
Fourth Lactations		
December-February	0.87504	0.81749
March-May	0.84229	0.81719
June-August	0.86304	0.85181
September-November	0.87010	0.82145
Fifth Lactations		
December-February	0.87229	0.83309
March-May	0.85258	0.83257
June-August	0.87151	0.84175
September-November	0.87546	0.82620
Sixth Lactations and over		
December-February	0.87002	0.81786
March-May	0.89227	0.84888
June-August	0.90886	0.89634
September-November	0.90143	0.85432
	Average Per Cent Fat-3.5% to 4.0%	
First Lactations		
December-February	0.94415	0.91706
March-May	0.92286	0.91704
June-August	0.93070	0.94650
September-November	0.95662	0.93505

Table 37

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 3

[Colorado]

	Average Per Cent Fat-3.5% to 4.0%	
	Milk	Fat
Second Lactations		
December-February	0.89992	0.86836
March-May	0.86982	0.85799
June-August	0.85747	0.86237
September-November	0.89409	0.87832
Third Lactations		
December-February	0.87155	0.82815
March-May	0.86726	0.85267
June-August	0.84184	0.83772
September-November	0.87047	0.83296
Fourth Lactations		
December-February	0.90326	0.85692
March-May	0.87015	0.84516
June-August	0.84392	0.83993
September-November	0.85579	0.83736
Fifth Lactations		
December-February	0.89150	0.85435
March-May	0.88049	0.85952
June-August	0.84852	0.84518
September-November	0.87878	0.84243
Sixth Lactations and over		
December-February	0.88858	0.85327
March-May	0.86636	0.87237
June-August	0.86814	0.85770
September-November	0.88790	0.84342
	Average Per Cent Fat-4.1% to 4.6%	
First Lactations		
December-February	0.94226	0.93816
March-May	0.90472	0.92676
June-August	0.87944	0.91615
September-November	0.93037	0.94443
Second Lactations		
December-February	0.89928	0.89063
March-May	0.86949	0.88939
June-August	0.84575	0.88978
September-November	0.88078	0.88362

Table 37

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 3

[Colorado]

	Average Per Cent Fat-4.1% to 4.6%	
	Milk	Fat
Third Lactations		
December-February	0.88063	0.85789
March-May	0.82383	0.82825
June-August	0.82909	0.87389
September-November	0.87916	0.86218
Fourth Lactations		
December-February	0.88627	0.86109
March-May	0.84186	0.86211
June-August	0.79914	0.79037
September-November	0.85912	0.84817
Fifth Lactations		
December-February	0.87543	0.84885
March-May	0.81644	0.83177
June-August	0.80956	0.86313
September-November	0.88040	0.87748
Sixth Lactations and over		
December-February	0.88537	0.85081
March-May	0.84590	0.85953
June-August	0.85201	0.86224
September-November	0.89990	0.89454
	Average Per Cent Fat-4.7% to 5.2%	
First Lactations		
December-February	0.94903	0.93624
March-May	0.87499	0.89358
June-August	0.88451	0.95502
September-November	0.91333	0.94976
Second Lactations		
December-February	0.87136	0.86134
March-May	0.84463	0.87035
June-August	0.83783	0.87729
September-November	0.86365	0.89091
Third Lactations		
December-February	0.87741	0.85538
March-May	0.83978	0.83574
June-August	0.77404	0.82743
September-November	0.87350	0.88111

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 3
[Colorado]

	Average Per Cent Fat-4.7% to 5.2%	
Fourth Lactations	Milk	Fat
December-February	0.88909	0.88080
March-May	0.86048	0.86232
June-August	0.84025	0.91914
September-November	0.84427	0.85796
Fifth Lactations		
December-February	0.86703	0.83433
March-May	0.82014	0.86371
June-August	0.79757	0.81305
September-November	0.82974	0.82406
Sixth Lactations and over		
December-February	0.87770	0.87714
March-May	0.83520	0.83091
June-August	0.88240	0.91964
September-November	0.82938	0.83946
	Average Per Cent Fat-5.3% and over	
First Lactations		
December-February	0.90768	0.90295
March-May	0.85092	0.85176
June-August	0.80793	0.91931
September-November	0.87951	0.90337
Second Lactations		
December-February	0.87896	0.86646
March-May	0.82964	0.83087
June-August	0.79667	0.85690
September-November	0.85215	0.85429
Third Lactations		
December-February	0.85910	0.84919
March-May	0.78826	0.80991
June-August	0.80436	0.86272
September-November	0.83255	0.84949
Fourth Lactations		
December-February	0.80930	0.81816
March-May	0.85080	0.89435
June-August	0.74595	0.74782
September-November	0.83255	0.83553

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 3
[Colorado]

	Average Per Cent Fat-5.3% and over	
Fifth Lactations	Milk	Fat
December-February	0.81737	0.79804
March-May	0.84924	0.89982
June-August	0.81101	0.84199
September-November	0.83645	0.80185
Sixth Lactations and over		
December-February	0.87260	0.86680
March-May	0.84191	0.83981
June-August	0.85028	0.84056
September-November	0.90363	0.88162

Area 5
[Utah and Nevada]

	Average Per Cent Fat-3.4% and under	
First Lactations		
December-February	0.94848	0.91579
March-May	0.92042	0.91416
June-August	0.91874	0.93013
September-November	0.94746	0.92504
Second Lactations		
December-February	0.87976	0.84133
March-May	0.85985	0.84084
June-August	0.85817	0.85834
September-November	0.88766	0.87006
Third Lactations		
December-February	0.87075	0.81994
March-May	0.84668	0.81092
June-August	0.83825	0.83417
September-November	0.87346	0.84741
Fourth Lactations		
December-February	0.87173	0.82786
March-May	0.86776	0.82983
June-August	0.85893	0.85154
September-November	0.88121	0.84049

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 5
[Utah and Nevada]

	Average Per Cent Fat-3.4% and under	
Fifth Lactations	Milk	Fat
December-February	0.88603	0.82280
March-May	0.86094	0.82673
June-August	0.96444	0.84747
September-November	0.87590	0.84083
Sixth Lactations and over		
December-February	0.88036	0.81812
March-May	0.87290	0.83547
June-August	0.85553	0.83617
September-November	0.88491	0.85101
	Average Per Cent Fat-3.5% to 4.0%	
First Lactations		
December-February	0.93392	0.90651
March-May	0.91176	0.91519
June-August	0.91057	0.93349
September-November	0.94253	0.93186
Second Lactations		
December-February	0.87608	0.84104
March-May	0.85489	0.83851
June-August	0.84995	0.85913
September-November	0.88275	0.85687
Third Lactations		
December-February	0.87879	0.82594
March-May	0.85307	0.82043
June-August	0.83431	0.84498
September-November	0.87418	0.83503
Fourth Lactations		
December-February	0.88076	0.82277
March-May	0.85698	0.82773
June-August	0.84185	0.83964
September-November	0.84895	0.80410
Fifth Lactations		
December-February	0.88043	0.80872
March-May	0.85583	0.82399
June-August	0.85240	0.84628
September-November	0.87293	0.82838

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 5
[Utah and Nevada]

	Average Per Cent Fat-3.5% to 4.0%	
Sixth Lactations and over	Milk	Fat
December-February	0.90037	0.84609
March-May	0.88065	0.84674
June-August	0.84943	0.85522
September-November	0.87188	0.83290
	Average Per Cent Fat-4.1% to 4.6%	
First Lactations		
December-February	0.90945	0.89810
March-May	0.89863	0.92431
June-August	0.88099	0.90958
September-November	0.92821	0.92359
Second Lactations		
December-February	0.87754	0.84885
March-May	0.83577	0.79908
June-August	0.84643	0.87093
September-November	0.86259	0.85513
Third Lactations		
December-February	0.86541	0.82389
March-May	0.83577	0.79908
June-August	0.84821	0.83405
September-November	0.86070	0.84585
Fourth Lactations		
December-February	0.86970	0.83887
March-May	0.82685	0.83627
June-August	0.83768	0.85574
September-November	0.86340	0.84150
Fifth Lactations		
December-February	0.87681	0.83712
March-May	0.84352	0.86281
June-August	0.77762	0.83128
September-November	0.87267	0.84389
Sixth Lactations and over		
December-February	0.88846	0.84888
March-May	0.82857	0.84060
June-August	0.83081	0.83756
September-November	0.88255	0.84771

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 5

[Utah and Nevada]

	Average Per Cent Fat-4.7% to 5.2%	
	Milk	Fat
First Lactations		
December-February	0.89898	0.92045
March-May	0.89249	0.95497
June-August	0.84644	0.92848
September-November	0.90990	0.94942
Second Lactations		
December-February	0.83885	0.84336
March-May	0.82222	0.85068
June-August	0.81029	0.87202
September-November	0.84253	0.86131
Third Lactations		
December-February	0.87432	0.86090
March-May	0.81367	0.83927
June-August	0.80241	0.86298
September-November	0.82532	0.82489
Fourth Lactations		
December-February	0.86611	0.86721
March-May	0.80688	0.83264
June-August	0.81913	0.85288
September-November	0.87611	0.86729
Fifth Lactations		
December-February	0.86042	0.84438
March-May	0.86232	0.88174
June-August	0.86235	0.91196
September-November	0.85532	0.84073
Sixth Lactations and over		
December-February	0.89754	0.88528
March-May	0.86439	0.86323
June-August	0.76302	0.81346
September-November	0.85547	0.86104
	Average Per Cent Fat-5.3% and over	
First Lactations		
December-February	0.89870	0.91197
March-May	0.89983	0.94857
June-August	0.86601	0.94769
September-November	0.85532	0.92300

Table 37
[continued]

Factors for use in estimating total from part milk and fat production in areas 3 and 5.

Area 5
[Utah and Nevada]

	Average Per Cent Fat-5.3% and over	
	Milk	Fat
Second Lactations		
December-February	0.85561	0.86477
March-May	0.83225	0.86051
June-August	0.81440	0.87190
September-November	0.85756	0.87734
Third Lactations		
December-February	0.88712	0.88287
March-May	0.82792	0.86089
June-August	0.81517	0.88322
September-November	0.83387	0.83992
Fourth Lactations		
December-February	0.87062	0.84704
March-May	0.80378	0.80724
June-August	0.78524	0.84819
September-November	0.83201	0.84819
Fifth Lactations		
December-February	0.88325	0.85886
March-May	0.80934	0.83069
June-August	0.77250	0.85461
September-November	0.86966	0.87328
Sixth Lactations and over		
December-February	0.90387	0.87715
March-May	0.85095	0.85654
June-August	0.83868	0.88072
September-November	0.90954	0.91289

3.5.4. Factors for estimating total from part production for milk and fat in terms of level of milk and level of fat production. The factors here presented for use in estimating total from part lactation production in terms of level of milk and fat production, although the effects are significant, are probably of more academic than practical interest. The difficulty in finding a

base for level of production when the lactation is incomplete is but one of the problems. For this particular study, these variables were considered separately. This makes it difficult to use them in connection with the other significant variables. It was nevertheless felt to be beneficial to include factors for these effects in a table to make them accessible. Tables 38 and 39 present the factors for these variables.

Table 38

Factors for use in estimating total from part production for both milk and fat production for level of milk production. [in pounds.]

	Milk Coefficient	Fat Coefficient
5,999 and under	0.83693	0.87885
6,000 to 6,999	0.84723	0.87871
7,000 to 7,999	0.86040	0.88212
8,000 to 8,999	0.87060	0.88320
9,000 to 9,999	0.87758	0.87932
10,000 to 10,999	0.88215	0.87966
11,000 to 11,999	0.88772	0.87994
12,000 to 12,999	0.88582	0.87153
13,000 to 13,999	0.88377	0.86681
14,000 to 14,999	0.88540	0.85751
15,000 to 15,999	0.88746	0.86236
16,000 to 16,999	0.88401	0.85458
17,000 to 17,999	0.89618	0.86129
18,000 to 18,999	0.89121	0.85755
19,000 to 19,999	0.90211	0.85814
20,000 and over	0.90526	0.86579

Table 39

Factors for use in estimating total from part production for both milk and fat production for level of fat production [in pounds].

	Milk Coefficient	Fat Coefficient
309 and under	0.86767	0.88294
310 to 339	0.87744	0.88138
340 to 369	0.87525	0.87609
370 to 399	0.88173	0.88246
400 to 429	0.88290	0.87532
430 to 459	0.88201	0.87323
460 to 489	0.88113	0.86874
490 to 519	0.88299	0.86582
520 to 549	0.88201	0.86192
550 to 579	0.88549	0.86269
580 to 609	0.89022	0.86355
610 to 639	0.88916	0.85955
640 to 669	0.88984	0.85802
670 to 699	0.90022	0.86603
700 and over	0.90060	0.85730

3.6. Summary of the discussion of the results. The analyses of variance indicated that all variables studied had an effect on the lactation estimation curve, either directly, or through interaction with other variables, with the exception of Days Dry and Days Open. The significant interactions indicated that Per Cent Fat interacted with climate either by interacting with Season of Freshening or by interacting with Geographic Area, although in the latter case, the interaction could be due to breed differences. An analysis of the components of variance indicated that Lactation Number accounted for the largest proportion of the variation, in one case accounting for more than 90%. Per Cent Fat and Season of Freshening accounted for relatively more variation for milk coefficients than for fat coefficients. An analysis of means indicated that certain

variables could be combined, provided they were involved in no interaction effects. Graphs of the coefficients indicated that coefficients for fat reacted similarly in both sets of areas, as did those for milk. While both sets of coefficients for milk and fat for Lactation Number in areas 2 and 4 upheld the previously discussed results that first lactations should be separated from the succeeding ones and that the succeeding ones could be estimated together, in areas 3 and 5, it was found that the second and sixth or greater lactations should also be estimated by a separate coefficient, although in the case of milk, separate estimation of the second lactation from the third, fourth, and fifth, does not seem to be quite so justified. It was found that a different combination of seasons should be used for milk than for fat with December through May and June through November being used for fat and September through February and March through August being used for fat.

A cubic trend was found to exist for both Level of Milk Production and Level of Fat Production, for both milk and fat coefficients. The trend was in an upward direction for milk coefficients for both level of milk and level of fat production, and in a downward direction for the fat coefficients in both cases. The effect was not as pronounced for Level of Fat Production as for Level of Milk Production.

Graphs of the significant interactions showed the effects of moving across all levels of one of the variables while holding the other variable constant for each of its levels. It was interesting to note the effects of Per Cent Fat with respect

to interaction. Evidently there was a "muddying" effect of Breed on Per Cent Fat. Per Cent Fat interacted almost consistently with season effects [the exception being for fat in areas 3 and 5, where there was quite a large Per Cent Fat by Area interaction]. Because of the significant interaction effects, which influenced the results of the mean comparison tests, one of the variables was combined with respect to level, but the coefficients were left in terms of cell means.

From all analyses, it was seen that the structure of the variables for milk was distinctly different than that for fat. The structural differences between areas 2 and 4 and areas 3 and 5 could possibly have been due to the differences in which the levels were constructed for Per Cent Fat for the two sets of areas, and also to the smaller numbers of observations involved in areas 2 and 4.

If one were to use one set of factors common to the western states, the factors for Utah and Nevada would probably be the best to use, as they are based on the largest number of observations, and because the majority of cows are of the Holstein breed.

As was stated before, the factors for level of production are more of an academic than practical interest because of the fact that in their present form, they cannot be used in connection with the other variables.

ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance and contributions of the many people who helped to make this thesis a reality. He expresses his sincere thanks to his committee, who helped to get the project under way; to Dr. Stanley Wearden of the Statistics Department at Kansas State University, his major professor, for his suggestions on how to begin and where to find the pertinent literature; to Dr. Gary Krause, of the Statistics Department, for his interest and consideration in taking time out from his busy schedule to listen to and evaluate the author's ideas on the statistical problems involved. He especially wishes to thank Dr. Keith Huston, of the Department of Dairy and Poultry Science, for the many hours that he was willing to spend in helping the author to gain the necessary background in dairy husbandry for this undertaking, for his willingness to listen to the author's ideas and to help him evaluate the many articles he came across in the literature, and for his interest and enthusiasm in this lengthy and labourious project.

The author also wishes to acknowledge the able assistance of G. Brian Bone, Systems Engineer with International Business Machines Corporation in Salt Lake City, Utah, and to thank him for being ready at the right time with the right ideas and answers to some of the thorny programming problems faced by the author on an unfamiliar machine. He also wishes to acknowledge the programming assistance of William Austin and Waine Ivie on

the IBM 7040 at Brigham Young University and the assistance of his brother, B. Lynn Crandall, for setting up certain of the programs on the IBM 1401/1311 system at the DHI Computing Service.

The author also especially wishes to acknowledge the assistance of his father, Bliss H. Crandall, for suggesting the problem to be studied and for specifying so well the problems involved. He also wishes to thank him for his many suggestions as the study was in progress.

Thanks should also go to Dr. Gary Carlson and Edwin Dean, of the Computer Research Center at Brigham Young University for making available the many hours of computer time on the IBM 7040 which were necessary for the completion of many analyses used in the thesis.

The writer would also like to acknowledge the computer time made available on the IBM 1401/1311 system at the DHI Computing Service and the IBM 7070/7094 system at the University of Washington, Seattle, Washington.

The writer wishes to thank Robert Appleman, Extension Dairyman at the University of California, for his interest in this thesis and his many helpful suggestions.

The author owes a special debt of gratitude to Dr. D. G. Chapman of the Mathematics Department at the University of Washington for his help in discussing and making precise many of the statistical ideas and techniques used in this study, for his reading of the statistical portions of the thesis and his helpful suggestions.

The author also wishes to acknowledge the many individuals, too numerous to name, classmates, professors, and others who have contributed in some small but important way to the development of this thesis.

REFERENCES

Anderson, R. L., and Bancroft, T. A.
Statistical theory in research. New York: McGraw-Hill, 1952.

Bereskin, B., and Freeman, A. E.
Genetic and environmental factors in dairy sire evaluation. I. Effects of herds, months, and year-seasons on variance among lactation records; repeatability and heritability. Journal of Dairy Science, March, 1965, 48: 347-51.

Brandt, A. E.
The analysis of variance in a 2 x 5 table with disproportionate frequencies. Journal of the American Statistical Association, 1933, 28: 164. (Cited in Yates, 1934).

Brownlee, K. A.
Statistical theory and methodology in science and engineering. New York: John Wiley and Sons, 1960. 570 p.

Bush, N., and Anderson, R. L.
A comparison of three different procedures for estimating variance components. Technometrics, November, 1963, 5: 421-440.

Cannon, C. Y., *et al.*
Predicting 305-day yields from short-time records. Journal of Dairy Science, December, 1942, 25: 991.

Church, Alonzo, Jr.
Analysis of data when the response is a curve. Technometrics, May, 1966, 8: 229-246.

Corley, E. L., and Jensen, E. L.
Monthly variation in DHIA test day milk fat production. Paper presented at the 60th annual meeting of the American Dairy Science Association, June 21, 1965, Lexington, Kentucky. [Abstract published in Journal of Dairy Science, June, 1965, 48: 826.]

Crandall, Vern J.
Unpublished research. DHI Computing Service, Provo, Utah. 1963.

Dairy-Herd-Improvement Letter, ARS-44-152. Vol. 40, No. 8.

Deming, W. E.
Some theory of sampling. New York: John Wiley and Sons, 1950.

Deming, W. E.
Sample design in business research. New York: John Wiley and Sons, 1960.

Dixon, W. J., and Massey, F. J., Jr.
Introduction to statistical analysis. New York: McGraw-Hill, 1951.

Dixon, W. J.
BMD Biomedical computer programs. Health Sciences Computing Facility, Department of Preventive Medicine, and Public Health, School of Medicine, University of California, Los Angeles, September 1, 1965.

Federer, W. T.
Experimental design. New York: Macmillan, 1955. 544 p.

Fieller, E. C.
The distribution of the index in a normal bivariate population. Biometrika, 1932, 24: 428-440.

Fieller, E. C.
The biological standardization of insulin. Journal of the Royal Statistical Society, 1940, supplement, 7, 1-64. [Cited in Finney, 1964, and Deming, 1950].

Finney, D. J.
Statistical methods in bioassay. London: Charles Griffin, 1964. 668 p.

Gaines, W. L.
Deferred short-time test as a measure of the production of dairy cows. Journal of Agricultural Research, 1927, 35: 237-49. [Cited in Cannon, _et_. _al_., 1942].

Geary, R. C.
The frequency distribution of the quotient of two normal variates. Journal of the Royal Statistical Society, 1930, 93: 442-446.

Gosslee, D. G., and Lucas, H. L.
Analysis of variance of disproportionate data when interaction is present. Biometrics, March, 1965, 21: 115-133.

Goulden, C. H.
Methods of statistical analysis, New York: John Wiley and Sons, 1952. 467 p.

Graybill, Franklin A.
An introduction to linear statistical models, vol. 1. New York: NcGraw-Hill, 1961. 463 p.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G.
Sample survey methods and theory, Vol. 1. New York: John Wiley and Sons, 1956.

Harvey, W. R.
Least-squares analysis of data with unequal subclass numbers. ARS-20-8, July, 1960, U. S. Department of Agriculture.

Heady, E. O., and Dillon, J. L.
Agricultural production functions. Iowa State University Press, 1961.

Heady, E. O., Madden, J. P., Jacobson, N. L., and Freeman, A. E.
Milk production functions incorporating variables for cow characteristics and environments. Journal of Farm Economics, 1964, 46: 1-19.

Henderson, C. R.
Estimation of variances and covariance components. Biometrics, June, 1953, 9: 226-52.

Hickman, C. G., and Henderson, C. R.
Components of the relationship between level of production and rate of maturity in dairy cattle. Journal of Dairy Science, August, 1955, 38: 883-890.

Hickman, C. G.
Effects of level of herd environment. I. Relationship between yield and age. Journal of Dairy Science, July 1962, 45: 861-4.

Johansson, Ivar
Genetic aspects of dairy cattle breeding. Urbana, Illinois: University of Illinois Press, 1961.

Johansson, Ivar
The relation between body size, conformation, and milk yield in dairy cattle. Animal Breeding Abstracts, December, 1964, 32: 421-435.

Kartha, K. P. A.
A study of the data of the milk yields of the various types of cattle obtained from records of the government military dairy farms. Indian Journal of Veterinary Science and Animal Husbandry, 1934, 4: 218-231. [Cited in Cannon, et al., 1942.]

Kendall, M. G.
Advanced theory of statistics, vol. 1. [New Edition]. London: Charles Griffin, 1958.

Kendrick, J. F.
Standardizing Dairy-Herd Improvement-Association records in proving sires. United States Department of Agriculture ARS-52-1, January, 1955.

King, Gerald J., and Miller, Robert H.
The National Cooperative Dairy Herd Improvement Program handbook. Agricultural Handbook No. 278, Agricultural Research Service, U. S. Department of Agriculture, April, 1965.

Kramer, C. Y.
On the analysis of variance of a two-way classification with the unequal subclass numbers. Biometrics, 1955, 11: 441-52.

Lamb, R. C., and McGilliard, L. D. [a]
Variables affecting ratio factors for estimating 305-day production from part lactations. Journal of Dairy Science, April, 1960, 43: 519-28.

Lamb, R. C., and McGilliard, L. D. [b]
Comparison of ratio factors for extending part-time milk and fat records. Journal of Dairy Science, 1960, 43: 879.

Leroy, H. L., and Gluckowski, W.
Die bestimmung der varianz komponenten im a b c-factoren versuch mit ungleichen klassen frequenzen. Biometrische Zeitschrift, 1961, 3: 73-91. [Cited in Bush and Anderson, 1963].

Lush, J. L., and Shrode, R. R.
Changes in milk production with age and milking frequency. Journal of Dairy Science, May, 1950, 33: 338.

Madden, D. E., <u>et al</u>.
Relations between parts of lactations and producing ability of Holstein cows. Journal of Dairy Science, 1955, 38: 1264.

Madden, D. E., McGilliard, L. D., and Ralston, N. P.
Relations between monthly test day production of Holstein-Friesian cows. Journal of Dairy Science, 1956, 39: 932.

Madden, D. E., McGilliard, L. D., and Ralston, N. P.
Relations between test day milk production of Holstein dairy cows. Journal of Dairy Science, 1959, 42: 319.

Marsaglia, George
Ratios of normal variables and ratios of sums of uniform variables. Journal of the American Statistical Association, March, 1965, 60: 193-204.

Merrill, A. S.
Frequency distribution of an index when both the components follow the normal law. Biometrika, 1928, 20A: 53-63. [Cited in Fieller, 1932].

Mielke, Paul W., and McHugh, Richard B.
Non-orthogonality in the two-way classification for the mixed effects finite population model. Biometrics, June, 1965, 21: 308-323.

Miller, R. H.
Biases in the estimation of the regression of milk production on age. Journal of Dairy Science, August, 1964, 47: 855-60.

Ostle, B.
Statistics in research. Ames, Iowa: Iowa State College Press, 1954.

Rao, C. R.
Some statistical methods for comparison of growth curves. Biometrics, March 1958, 14: 1-17.

Searle, S. R., and Henderson, C. R.
Establishing age-correction factors related to the level of hard production. Journal of Dairy Science, May, 1959, 42: 824-835.

Searle, S. R.
Simplified herd-level age-correction factors. Journal of Dairy Science, June, 1960, 43: 821-24.

Searle, S. R., and Henderson, C. R.
Judging the effectiveness of age-correction factors. Journal of Dairy Science, July, 1960, 43: 966-74.

Searle, S. R. [a]
Part lactations. I. Age-corection factors for monthly milk fat yeidls. Journal of Dairy Science, January, 1961, 44:104-114.

Searle, S. R. [b]
Part lactations. II. Genetic and phenotypic studies of monthly milk fat yield. Journal of Dairy Science, February, 1961, 44: 282-95.

Searle, S. R. [c]
Part lactations. III. Progeny testing with part lactation records. Journal of Dairy Science, May, 1961, 44: 921-7.

Sheffé, Henry
The analysis of variance. New York: John Wiley and Sons, 1959. 477 p.

Smith, Vearle
Physiology of lactation. Ames, Iowa: Iowa State University Press, 1959.

Snedecor, G. W.
The method of expected numbers for tables of multiple classification with disproportionate subclass numbers. Journal of the American Statistical Association, 1934, 29: 389-93.

Snedecor, G. W.
Statistical methods. Ames, Iowa: Iowa State College Press, 1956. 534 p.

Steel, R. G. D., and Torrie, J. H.
Principles and procedures of statistics with special reference to the biological sciences. New York: McGraw-Hill, 1960.

Sundaresan, D., and Freeman, A. E.
Variation in milk fat production in state-owned herds in Iowa. Journal of Dairy Science, January, 1961, 44: 151-162.

Van Vleck, L. D., Heidhues, T., and Henderson, C. R.
Analysis of deviations of dairy records from different contemporary averages. Journal of Dairy Science, February, 1961, 44: 269-81.

Van Vleck, L. D., and Henderson, C. R. [a]
Estimates of genetic parameters of some functions of part lactation milk records. Journal of Dairy Science, June, 1961, 44: 1073-84.

Van Vleck, L. D., and Henderson, C. R. [b]
Regression factors for extending part lactation milk records. Journal of Dairy Science, June, 1961, 44: 1085-92.

Van Vleck, L. D., and Henderson, C. R. [c]
Ratio factors for adjusting monthly test-day data for age and season of calving and ratio factors for extending part lactation records. Journal of Dairy Science, June, 1961, 44: 1093-1102.

Van Vleck, L. D., and Henderson, C. R. [d]
Regression factors for predicting a succeeding complete lactation milk record from part lactation records. Journal of Dairy Science, July, 1961, 44: 1322-7.

Van Vleck, L. D., and Henderson, C. R. [e]
Extending part lactation milk records by regression ignoring herd effects. Journal of Dairy Science, August, 1961, 44: 1519-28.

Van Vleck, L. D., and Henderson, C. R. [f]
Variance and covariance components in part lactation milk and fat records. Journal of Dairy Science, October, 1961 44: 1780-6.

Wakefield, Gordon D.
Computer implementation of a non-orthogonal econometric analysis: estimation of the influences of selected socio-economic variables upon income. Unpublished master's thesis. Brigham Young University, Provo, Utah, 1965.

Wilks, S. S.
Mathematical Statistics. New York: John Wiley and Sons, 1962.

Yates, F.
The analysis of multiple classifications with unequal numbers in the different classes. Journal of the American Statistical Association, 1934, 29: 51-66.

APPENDIX

APPENDIX I

On the following pages will be found examples of the various forms of the data as it was used as the input and obtained as output for the principal programs and analyses described in the section on methods and materials. After each example will be found a description or discussion of the data.

FIGURE 24

DATA FORMAT AFTER LOADED ON TAPE AND SORTED INTO SEQUENCE BY YEAR AND MONTH WITHIN COW.

```
SORF OHIA RECORDS DHI COMP SER                                                        02/27/85 PAGE 2
                                                                        M = Code        6(80) From Code
       5.5U06   T86     0220800 FILE 1           RECORD  1        1200  (8)          640  '1101  WORDS
00000  8718108480182102763100001 30A00910  43A37  043C0902 00+05+00+15C070E0025F     36150404041704
00020  8718100480102102763100001 30A00910  43+33  041C1002 09+05+00+18C080100296     46150504051805
00040  8718106480182102763100001 30A00910  43A34  030A080212405+00+2101004003 3E     50150004001006
00060  8718108480182102763100001 30A00910  43+32  023H070215405+00+240107800035H     66150704072207
00100  871810648018410066220001 205A0             2H30   054H0901000000 00+14B067100210      26150363031603   Lactation in Progress 142 Days
00120  871810648018410066220001 205A0             3A30   050+100111E00000+17C083000204      36150403041304
00140  871810640018410066220001 205A0             3+30   045B090114E00000+20C097+00308      46150503051805
00160  871810648018410066220001 205A0             3A31   045H080117F00000+23D111800034F     56150663061506
00200  871810648018410066220001 205A0             3+30   041H090120F00000+26D123G00300      66150763071207
00220  871810648018410066220001 205A0             3A30   038H080123G00000+29E135G0042+      76150863081708   Lactation Completed At 301 Days. Note "R".
00240  8718100480184 0805   D-Y   0001205A0 1006622  DF30  038H000124C0000KE30A130+R042G     86150903091509
00260  871810648018409046310 00        0805       42C30   052H1502   03800+02C012A00046     96151083101910   New Lactation Begins 9/04/63
00300  871810648018400046310 00        0805       43A30   073F1202   03800+050034100 11E    06151183112312
00320  8718106480184090463100011 16A00805         43+34   071F130201A03800+080050D0010H     N6151263121412
00340  8718106480184090463100011 16A00805         43A31   086C150204803800+11E077+00250     D6150104011801
00300  87181064001040904631 0000106A30805         43A30   046F120202D03800+14F091D0029E     18150204022202
00400  87181064801040904031 0000106A30805         42131   046E000205C03800+17E104100336     26150364031403
00420  87181004801840904831 0000106A30805         43A33   042+090208003800+20F117100 30+    36150404041704
00440  87181004801840904831 0000106A30805         43+31   030E080211003800+23F120100410     46150564051605
00460  87181004801840904631 0000106A30805         43A32   032H070214E03800+26G1394 0044G    56150664061600   Lactation Still in Progress at 297 Days
00500  87181084801840904631 0000106A30805         43+30   018C060217E03000+29G144F00400     60150764072207
00520  871810646016510016220001 221A0             2H31   036G080100H00000+14G055100 10F     28150363031003
00540  871810640018510016220001 221A0             3A30   034007010910 0000+17M000E0021H     30150463041304
00580  871810648018510016220001 221AD             3+31   035+000112100000+20M077+00254      48150563051005
00000  871810848018510018220001 22 AD             3A32   029C070116+00000+23100640020+      50150603001500
00620  871810848018510016220001 221A0             3+31   034C060119+00000+26109600 0310     46150763071207
```

IDENTIFICATION	Current Status Date	Body Wt	Breeding Date	Previous Status	This Period	LACTATION To Date	Dates For Following Month
Herd, Cow, ...	Mo, Day, Yr		Mo, Day, ...	Mo, Day, Yr	Daily Milk, % Fat, % SNF, % Prot ...	Calv, Days: Mo Day, Days Dry, In Milk, Pounds Milk (×10), Pounds Fat, Pounds SNF, Lbs Prot	Center Day, Tested, ...

17 Months Records for Cow 18K

Figure 24 shows the form of the data after they had been loaded on tape and sorted into sequence by year and month within cow. The record for the cow which is marked represents 17 monthly production records beginning when the cow was 142 days into her lactation. The "R" is the flag mentioned in the section on methods which indicates that either the lactation is complete or that the cow has milked one month past 305 days. For this particular case, the first part of the lactation is missing, and thus, this part of the cow's record would not be used. Also, the lactation was completed at 301 days, hence, if the early part were present, the record would be a "short lactation" and still could not be used. The second lactation for this cow also cannot be used, as the latter portion of the lactation is not available. This is the basic type of record from which the data were taken. These records were the input to the edit program.

FIGURE 25

DATA FORMAT FOR EDITED RECORDS.

Figure 25 is the output to the edit program and the input to the regression setup program [see Figure 26]. There is a great deal of information in these data. As can be seen the record begins with the information necessary to set up the variable codes for the 8 variables to be studied. Days in milk has been rolled back to 305 days and a 305-Day record has been computed. The average per cent fat is included as a part of the 305-Day record. Following this is given the per cent fat and pounds of milk produced for the current period [the milk is expressed to nearest 10 pounds] and the cumulative days in milk, pounds of milk, and pounds for that period for eight to eleven consecutive periods. [The cumulative milk is expressed to the nearest 10 pounds, and the cumulative fat to the nearest pound.] Also included is the centering day and the test day for each of the first 10 periods.

FIGURE 26

DATA FORMAT FOR REGRESSION SETUP TAPE.

Computer Code	Variable Code	Days Milk	305-Day Milk	Cum. Milk	305-Day Fat	Cum. Fat	Period Number	No. Reports		Cow Number
CC0484	1312190503	013	0989	0047	0346	0014	01	11	1	000001
CC0484	1312190503	044	0989	0165	0346	0052	02	11	1	000001
CC0484	1312190503	074	0989	0281	0346	0089	03	11	1	000001
CC0484	1312190503	105	0989	0367	0346	0117	04	11	1	000001
CC0484	1312190503	136	0989	0461	0346	0150	05	11	1	000001
CC0484	1312190503	166	0989	0566	0346	0185	06	11	1	000001
CC0484	1312190503	197	0989	0659	0346	0219	07	11	1	000001
CC0484	1312190503	227	0989	0759	0346	0252	08	11	1	000001
CC0484	1312190503	258	0989	0844	0346	0285	09	11	1	000001
CC0484	1312190503	289	0989	0943	0346	0326	10	11	1	000001
CC0484	1312190503	318	0989	1021	0346	0357	11	11	1	000001
CC0484	1224450704	031	1145	0186	0389	0060	01	11	1	000002
CC0484	1224450704	062	1145	0322	0389	0128	02	11	1	000002
CC0484	1224450704	090	1145	0454	0389	0176	03	11	1	000002
CC0484	1224450704	121	1145	0572	0389	0218	04	11	1	000002
CC0484	1224450704	151	1145	0676	0389	0249	05	11	1	000002
CC0484	1224450704	182	1145	0783	0389	0281	06	11	1	000002
CC0484	1224450704	212	1145	0887	0389	0305	07	11	1	000002
CC0484	1224450704	243	1145	0992	0389	0337	08	11	1	000002
CC0484	1224450704	274	1145	1077	0389	0365	09	11	1	000002
CC0484	1224450704	304	1145	1143	0389	0389	10	11	1	000002
CC0484	1224450704	335	1145	1196	0389	0405	11	11	1	000002
CC0484	1313130706	030	1198	0164	0439	0039	01	11	1	000003
CC0484	1313130706	061	1198	0317	0439	0088	02	11	1	000003
CC0484	1313130706	091	1198	0449	0439	0129	03	11	1	000003
CC0484	1313130706	122	1198	0571	0439	0172	04	11	1	000003
CC0484	1313130706	152	1198	0675	0439	0214	05	11	1	000003
CC0484	1313130706	183	1198	0776	0439	0260	06	11	1	000003
CC0484	1313130706	214	1198	0903	0439	0321	07	11	1	000003
CC0484	1313130706	243	1198	1006	0439	0357	08	11	1	000003
CC0484	1313130706	274	1198	1110	0439	0404	09	11	1	000003
CC0484	1313130706	304	1198	1196	0439	0438	10	11	1	000003
CC0484	1313130706	335	1198	1278	0439	0472	11	11	1	000003
CC0484	1311120705	045	1174	0194	0425	0076	01	10	1	000004
CC0484	1311120705	076	1174	0343	0425	0133	02	10	1	000004
CC0484	1311120705	106	1174	0502	0425	0181	03	10	1	000004
CC0484	1311120705	137	1174	0624	0425	0215	04	10	1	000004
CC0484	1311120705	168	1174	0767	0425	0265	05	10	1	000004
CC0484	1311120705	198	1174	0878	0425	0298	06	10	1	000004
CC0484	1311120705	229	1174	0983	0425	0337	07	10	1	000004
CC0484	1311120705	259	1174	1064	0425	0374	00	10	1	000004
CC0484	1311120705	290	1174	1137	0425	0409	09	10	1	000004
CC0484	1311120705	321	1174	1213	0425	0442	10	10	1	000004
CC0484	1241550804	041	1256	0254	0397	0084	01	11	1	000005
CC0484	1241550804	072	1256	0394	0397	0141	02	11	1	000005
CC0484	1241550804	102	1256	0538	0397	0183	03	11	1	000005
CC0484	1241550804	133	1256	0705	0397	0238	04	11	1	000005
CC0484	1241550804	163	1256	0834	0397	0257	05	11	1	000005
CC0484	1241550804	194	1256	0975	0397	0305	06	11	1	000005
CC0484	1241550804	225	1256	1074	0397	0334	07	11	1	000005
CC0484	1241550804	255	1256	1163	0397	0366	08	11	1	000005
CC0484	1241550804	286	1256	1231	0397	0386	09	11	1	000005
CC0484	1241550804	316	1256	1270	0397	0402	10	11	1	000005
CC0484	1241550804	315	1256	1269	0397	0402	11	11	1	000005
CC0484	1311160402	030	0874	0093	0311	0052	01	11	1	000006
CC0484	1311160402	058	0874	0208	0311	0093	02	11	1	000006
CC0484	1311160402	089	0874	0304	0311	0127	03	11	1	000006
CC0484	1311160402	119	0874	0396	0311	0157	04	11	1	000006
CC0484	1311160402	150	0874	0487	0311	0184	05	11	1	000006
CC0484	1311160402	180	0874	0579	0311	0215	06	11	1	000006
CC0484	1311160402	211	0874	0663	0311	0234	07	11	1	000006
CC0484	1311160402	242	0874	0739	0311	0264	08	11	1	000006
CC0484	1311160402	272	0874	0805	0311	0287	09	11	1	000006
CC0484	1311160402	303	0874	0870	0311	0310	10	11	1	000006

Figure 26 is the input to the regression program mentioned on page 34. The data were taken from the output to the Data Edit program and transformed into the format shown above. In doing so, the data were coded according to the 8 variables to be studied. These codes are found in the area labeled "Variable Code," and represent Geographic Area [1st column], Per Cent Fat Level [2nd column], Lactation Number [3rd column], Season [4th column], Previous Days Dry [5th column], Days open [6th column], Level of Milk Production [7-8th columns], and Level of Fat Production [9-10th columns]. The first area is a computer code designating the type of record and the number of tape positions in it. It is for machine use only. The third field or area is the cumulative days in milk to that particular period. The fourth field is total 305-Day milk production. The fifth is the cumulative milk production to the given period. The sixth is the total 305-Day fat production; the seventh is the cumulative fat production to the given period. The eighth is period for which the cumulative record is given. The ninth is the total number of periods in the lactation, or 305-Day record. The final field is the cow number, a number which was assigned in consecutive order to the cows used in the study. The majority of the 305-Day records used had 11 periods contained in them.

Note that Cow Number 1 had 318 days in her record. This was rolled back to 305 days, and the production was adjusted accordingly for both milk and fat to give the total 305-Day production. The complete record to 318 days was used in

FIGURE 27

DATA FORMAT FOR REGRESSION OUT PUT TAPE

Regression Coefficient Milk	Y-Intercept Milk	Regression Coefficient Fat	Y-Intercept Fat	Σx^2	$\Sigma x y_{milk}$	Σy^2_{milk}	$\Sigma X'$	$\Sigma Y'_{milk}$	$\Sigma x y_{fat}$	Σy^2_{fat}	$\Sigma Y'_{fat}$	$\Sigma d^2_{y.x(milk)}$	$\Sigma d^2_{y.x(fat)}$	No. Records	Cow No.	Variable Code
-0.95782	-0.01648	-0.93851	-0.02224	0.92944	-0.89023	0.85659	-3.16592	2.85107	-0.87229	0.82614	2.72661	0.00390	0.00749	11	1	1263301510
-0.89677	-0.01309	-0.88156	-0.00588	0.86706	-0.77755	0.70144	-3.01963	2.56393	-0.76437	0.67692	2.59725	0.00416	0.00308	11	2	1161600702
-1.06167	-0.03536	-0.95861	-0.01507	1.49665	-1.58894	1.69581	-4.07935	3.94199	-1.43471	1.37663	3.74476	0.00888	0.00130	11	3	1274500804
-0.80905	-0.01892	-0.86308	-0.02322	0.94655	-0.76581	0.62293	-3.30712	2.46752	-0.81694	0.70966	2.59885	0.00335	0.00457	11	4	1264400904
-0.95076	-0.01057	-0.80959	0.00436	0.86328	-0.82077	0.78754	-3.00670	2.74235	-0.69890	0.56643	2.48211	0.00718	0.00061	11	5	1231600805
-0.89225	-0.00151	-0.82391	0.01786	0.83433	-0.74443	0.66503	-2.94357	2.60980	-0.68741	0.57285	2.62174	0.00082	0.00648	11	6	1232200904
-0.81846	-0.01293	-0.80697	-0.02324	0.96666	-0.79117	0.64962	-3.27651	2.53952	-0.78007	0.63547	2.38846	0.00208	0.00598	11	7	1224400703
-0.90390	-0.01142	-0.89728	-0.00445	1.00270	-0.90635	0.82101	-3.35519	2.90711	-0.89970	0.80804	2.96159	0.00176	0.00075	11	8	1211100501
-0.95298	-0.00723	-0.99820	0.00469	1.74794	-1.66575	1.58957	-4.30040	4.01868	-1.74480	1.74355	4.34429	0.00215	0.00188	11	9	1312100503
-0.82176	-0.00201	-0.83902	0.00554	0.89912	-0.73886	0.60812	-3.09325	2.51983	-0.75430	0.63423	2.65621	0.00095	0.00124	11	10	1312100402
-0.01017	-0.00368	-0.07144	0.00482	0.81481	-0.66014	0.53633	-2.89118	2.30193	-0.71006	0.61986	2.57251	0.00151	0.00110	11	11	1212100703
-0.98487	-0.00149	-1.13079	0.00493	0.84446	-0.83168	0.81963	-2.97549	2.81412	-0.95491	1.08172	3.41892	0.00053	0.00191	11	12	1213100501
-0.79433	-0.00815	-0.76865	-0.02021	1.03185	-0.81963	0.65212	-3.40986	2.61890	-0.79313	0.62113	2.39867	0.00107	0.01148	11	13	1224400704
-0.84574	-0.00505	-1.03447	-0.00655	1.05592	-0.89303	0.75610	-3.41346	2.83120	-1.09232	1.13133	3.45909	0.00083	0.00136	11	14	1313100706
-1.00526	-0.00272	-0.94524	0.01033	1.32648	-1.33345	1.34091	-3.88137	3.87187	-1.25380	1.19010	3.78243	0.00045	0.00492	11	15	1114100401
-0.93181	-0.02961	-0.87826	-0.00980	0.67905	-0.63275	0.59526	-2.88293	2.39029	-0.59638	0.52566	2.43400	0.00566	0.00188	10	16	1311100705
-0.80858	-0.02089	-0.75766	-0.01674	0.82166	-0.66438	0.54129	-3.01165	2.20531	-0.62254	0.47542	2.09767	0.00409	0.00375	11	17	1241500804
-0.93503	-0.01990	-0.75722	-0.00620	1.07095	-1.00136	0.94268	-3.48309	3.03863	-0.81095	0.61501	2.56984	0.00638	0.00094	11	18	1311100402
-0.91050	-8.01407	-0.81583	-0.00677	0.90549	-0.82445	0.75252	-3.20263	2.76126	-0.73873	0.60507	2.53932	0.00186	0.00239	11	19	1211100603
-0.77441	-0.02975	-0.74492	-0.03194	1.01093	-0.78286	0.61226	-3.41729	2.31915	-0.75307	8.56771	2.19430	0.80599	0.00674	11	20	1331400909
-0.75045	-0.03093	-0.72912	-0.03141	0.89866	-0.74959	0.57355	-3.32661	2.15618	-0.72829	8.54251	2.08006	0.01102	0.01150	11	21	1341300908
-0.74859	-0.02098	-0.82949	-0.02288	0.83815	-0.82743	0.47391	-3.04813	2.05106	-0.69521	0.58201	2.27674	0.00422	0.00533	11	22	1443300505
-0.84202	-0.02418	-0.81404	-0.03007	1.41400	-1.19063	1.00772	-3.96641	3.07384	-1.15106	0.94332	2.89807	0.00518	0.00631	11	23	1424300303
-0.68898	-0.00569	-0.65984	-0.00143	1.15532	-0.79599	0.55208	-3.61219	2.42608	-0.76232	0.50690	2.36773	0.00366	0.00389	11	24	1344400402
-0.80871	-0.00039	-0.87771	0.00409	0.80233	-0.64885	0.52700	-2.88259	2.32689	-0.7042[illegible]	0.61960	2.57504	0.00227	0.00151	11	25	1263200401
-0.74394	-0.01749	-0.70331	-0.01781	1.07960	-0.80316	0.60078	-3.49700	2.40922	-0.75924	0.53723	2.26354	0.00327	0.00321	11	26	1361400504
-0.79376	-0.02936	-8.84590	-8.02602	1.22613	-0.97325	0.78055	-3.70703	2.61952	-1.03725	0.88362	2.04975	0.00801	0.00615	11	27	1243300502
-0.99447	-0.02252	-0.97027	-0.01960	1.33174	-1.32438	1.32060	-3.85893	3.58966	-1.29215	1.25611	3.52660	0.00354	0.00237	11	28	1324300806
-0.93915	0.00712	-0.96466	0.00339	1.02729	-8.96477	0.90810	-3.38028	3.25133	-0.99097	8.95805	3.30682	0.00211	0.00208	11	29	1414100201
-1.05979	-0.05002	-0.87179	-0.02330	1.19441	-1.30166	1.44345	-3.65912	3.43742	-1.04128	0.91155	2.93280	0.02490	0.00376	11	30	1324501010
-0.78057	-0.02588	-0.76496	-0.01508	1.25585	-0.98028	0.77867	-3.72405	2.62221	-0.96068	0.73820	2.68287	0.01349	0.00332	11	31	1323300604
-0.67482	-0.03266	-0.68076	-0.03072	0.67806	-0.45757	0.31668	-2.88688	1.62154	-0.46160	0.32184	1.65805	0.00791	0.00760	10	32	1331400503
-0.86701	-0.01534	-0.95613	-0.02013	0.81548	-0.70866	0.61879	-2.96089	2.40431	-0.77970	0.75086	2.60959	0.00295	0.00536	11	33	1423200302
-0.88670	-0.01792	-0.83626	-0.01612	0.61273	-0.54331	8.48411	-2.79712	2.30099	-0.51240	0.42993	2.17793	0.00236	0.00144	10	34	1464100606
-0.89732	-0.02918	-0.03709	-0.02411	1.05897	-0.95024	0.65233	-3.43155	2.76826	-0.88731	8.75078	2.61004	0.00965	0.00731	11	35	1464300505
-0.87243	-0.02313	-0.85775	-0.01942	0.61273	-0.53456	8.46827	-2.79712	2.20903	-0.52557	0.45201	2.20499	0.00190	0.00121	10	36	1434100403
-0.97074	0.00334	-1.04200	0.00679	0.93214	-0.90486	0.87873	-3.18266	3.12821	-0.97131	1.01311	3.39115	0.00035	0.00092	11	37	1443300606
-0.76267	-0.01689	-0.72710	-0.00927	0.86109	-0.65673	0.50524	-3.11079	2.18668	-0.82613	0.45683	2.15992	0.00437	0.00159	11	38	1344300909
-0.82722	-0.02783	-0.85109	-0.02707	1.19992	-0.99260	0.82854	-3.64914	2.71249	-1.02121	0.87719	2.80800	0.00744	0.00802	11	39	1412100404
-1.04969	0.00051	-1.11428	-0.00187	0.89268	-0.93704	0.98546	-3.08974	3.24883	-0.99473	1.11128	3.42225	0.00186	0.00291	11	40	1363500605
-0.73635	-0.02730	-0.71731	-0.02419	0.90062	-0.66318	0.49839	-3.09548	1.97901	-0.64602	0.47309	1.95437	0.01006	0.00969	11	41	1331200604
-8.74921	-0.00717	-0.77286	-0.81025	0.78222	-0.58605	0.44258	-2.84922	2.05577	-0.60455	0.47191	2.08929	0.00351	0.00466	11	42	1424500504
-0.90075	-0.02458	-0.84855	-0.02146	0.66561	-0.77970	0.70843	-3.06619	2.40145	-0.73452	0.62814	2.36573	0.00612	0.00487	11	43	1261500805
-0.01925	-0.02583	-0.86153	-0.02932	0.75959	-0.62229	8.51379	-3.11153	2.29085	-0.65441	0.56849	2.38750	0.00398	0.00470	10	44	1362400502
-0.82087	-0.01933	-0.76103	-0.01101	1.42072	-1.17362	0.96532	-4.02551	3.09177	-1.08921	0.83089	2.94563	0.00193	0.00109	11	45	1341400705
-0.82950	-0.02308	-0.83697	-0.02039	1.32654	-1.10036	0.92009	-3.87298	2.95877	-1.11027	0.93552	3.01726	0.00735	0.00627	11	46	1324300504
-0.81197	-0.01845	-0.82359	-0.01234	1.09147	-0.88625	0.72282	-3.49862	2.63780	-0.89893	0.74177	2.74571	0.00322	0.00143	11	47	1231301008
-0.93330	0.00313	-0.93062	0.00470	0.89306	-0.83349	0.77867	-3.12569	2.95159	-0.83111	0.77417	2.96052	0.00077	0.00073	11	48	1364400706
-0.95642	0.00352	-0.98123	0.00229	1.03109	-0.98616	0.94675	-3.40184	3.29229	-1.01174	0.99602	3.36318	0.00357	0.00327	11	49	1324200301
-0.83261	0.00646	-0.84957	0.00183	0.78350	-0.65236	0.54364	-2.97448	2.54763	-0.66564	0.56613	2.54720	0.00048	0.00062	11	50	1313100301
-0.80975	-0.00006	-0.79239	-0.00339	0.84680	-0.65569	8.55581	-2.99151	2.42174	-0.67099	0.53288	2.33311	0.00057	0.00119	11	51	1313100402
-0.88547	0.00750	-0.88011	0.00631	0.96723	-8.85645	0.75870	-3.24745	2.95799	-0.85901	0.76314	2.95355	0.80034	0.00024	31	52	1313100302
-0.98698	-0.03361	-0.89675	-0.02506	1.46563	-1.44655	1.43760	-4.01221	3.59023	-1.31431	1.18371	3.32226	0.80988	8.00510	11	53	1322400502
-0.99707	-0.01490	-0.92033	-0.00663	1.06019	-1.05708	1.05585	-3.52529	3.35014	-0.97572	0.89903	3.17154	0.00187	0.00106	11	54	1311100704
-1.01870	-0.01669	-0.96243	-0.01518	1.61329	-1.64345	1.68416	-4.18009	4.07466	-1.55268	1.50058	3.85602	0.00998	0.00624	11	55	1213100603
-0.98822	0.00200	-1.01635	-0.00901	1.01322	-1.00129	0.99012	-3.34570	3.32832	-1.02978	1.04743	3.30132	0.00062	0.00081	11	56	1213100804
-1.00306	-0.02304	-0.96053	-0.02580	1.55508	-1.55984	1.56932	-4.12028	3.87950	-1.49370	1.43912	3.67386	0.00470	0.00436	11	57	1314100806
-0.99542	-0.02573	-1.00745	-0.02847	1.82393	-1.81557	1.80946	-4.40187	4.09869	-1.83751	1.85470	4.12144	0.00221	0.00350	11	58	1214100702
-0.97688	-0.01048	-0.97052	-0.02503	0.82737	-0.80824	0.79295	-3.06652	2.79229	-0.81043	8.80036	2.72844	0.00340	0.00652	11	59	1314100502
-1.05716	-0.00196	-0.97373	-0.00454	0.94705	-1.03119	1.05951	-3.20745	3.36920	-0.92217	0.89840	3.07323	0.80109	0.00045	11	60	1214101007
-0.98454	-0.01613	-0.00844	-0.00129	1.19180	-1.17338	1.15803	-3.65640	3.40246	-0.95396	0.76525	2.91414	0.00278	0.00167	11	61	1311100800
-0.98947	-0.01755	-0.91823	-8.01252	0.79781	-0.78941	0.78555	-2.91947	2.89564	-0.73257	0.67486	2.54303	0.00445	0.00219	11	62	1211100603
-0.78273	-0.00583	-0.74729	0.00185	0.61998	-0.48528	0.38084	-2.74212	2.08801	-0.46331	0.34731	2.06772	0.00100	0.00109	10	63	1223500603
-0.74183	-0.02800	-0.75742	-0.02583	0.66463	-0.49304	0.37253	-2.87350	1.85167	-0.50340	0.38831	1.91818	0.00677	0.00702	10	64	1223400501
-0.86004	-0.00472	-0.81223	-0.01228	0.94784	-0.81518	0.70226	-3.24386	2.73796	-0.76957	0.62757	2.49969	0.00117	0.00226	11	65	1214100703
-0.97287	-0.02428	-0.92622	-0.01707	1.24249	-1.20878	1.17985	-3.75105	3.38223	-1.15082	1.06941	3.28653	0.00387	0.00350	11	66	1212100307
-0.80490	-0.01034	-0.96448	-0.01281	1.32656	-1.17387	1.04166	-3.84064	3.28486	-1.27944	1.23724	3.56328	0.00290	0.00324	11	67	1313100402
-0.83315	-0.03386	-0.80980	-0.03017	1.04482	-0.87050	0.73334	-3.48691	2.53272	-0.84609	0.69214	2.49186	0.00808	0.00696	11	68	1261400704
-0.98356	-0.01619	-8.03440	-0.01103	1.85519	-1.03784	1.02488	-3.45308	3.21823	-0.88045	0.73975	2.75989	0.00410	0.00510	11	69	1344300807
-0.82819	-0.01051	-0.92348	-0.03543	1.08741	-0.90058	0.74674	-3.56733	2.83886	-1.00420	0.94130	2.90461	0.00089	0.01394	11	70	1214100301
-0.91233	-0.02092	-0.87647	-0.01376	1.04057	-0.94341	0.87330	-3.37869	2.85255	-0.91203	0.80480	2.80998	0.00707	0.00544	11	71	1211100502
-0.84500	-0.02609	-0.74268	-0.01644	0.63297	-0.53486	0.45854	-2.75068	2.06343	-0.50174	0.40115	2.01598	0.00659	0.00343	10	72	1312100402
-0.88355	-0.01546	-0.86051	-0.01394	1.10461	-0.98261	0.87700	-3.53360	2.97328	-0.95052	0.81985	2.88730	0.00292	0.00192	11	73	1364101010
-0.90460	-0.01012	-0.78922	-0.00150	1.17913	-1.06664	0.96562	-3.65541	3.19534	-0.93059	0.73468	2.86838	0.00073	0.00024	11	74	1251501209
-0.88427	-0.01566	-0.88750	-0.01460	1.06195	-0.93905	0.83367	-3.48413	2.80865	-0.94248	0.83818	2.93151	0.00330	0.00173	11	75	1353301312
-0.89195	-0.00758	-0.86729	0.01282	1.03591	-0.92398	0.82561	-3.36152	2.91498	-0.89843	0.78060	3.05643	0.00146	0.00139	11	76	1242501613
-0.90668	-0.02376	-0.89788	-0.01009	1.36558	-1.23815	1.12628	-3.91265	3.28420	-1.22612	1.10152	3.40208	0.00367	0.00061	11	77	1232601207
-0.93707	-0.02696	-0.90624	-0.02027	0.91688	-0.85917	0.80974	-3.26566	2.76358	-0.83091	0.75548	2.73649	0.00483	0.00247	11	78	1334501211
-0.91594	0.00386	-0.84684	-0.00147	0.81738	-0.76502	0.71682	-2.89689	2.75377	-0.73306	0.65777	2.58191	0.00081	0.00033	11	79	1231601303
-0.85089	-0.01600	-0.93617	-0.01703	1.07936	-0.95081	0.84095	-3.45870	2.87074	-1.01047	0.94878	3.05058	0.00340	0.00282	11	80	1331301003
-0.90155	-0.02602	-0.85472	-0.01860	1.61143	-1.45279	1.31478	-4.22450	3.52239	-1.37732	1.17957	3.40620	0.00501	8.00235	11	81	1321400808

calculating the regression coefficients, sums of squares, etc., in the regression program.

Figure 27 gives the output to the regression program. Only the regression coefficients for milk and fat were used in the main analysis, however, the other values were included to retain information which could be used in later studies. The column headings should be self-explanatory, as the notation has been used elsewhere in the study. The Y-intercepts for milk and fat are merely $\log_{10}A$ for milk and fat, and the regression coefficients are the values of C in equation [6].

APPENDIX II

Over 50 programs were used to load the data on tape, list them, edit the records, convert them into workable form, etc. Most of these programs were of such a specialized nature that there would be little purpose in including them in this study. They were written for the IBM 1401/1311 system at the DHI Computing Service, the IBM 7040 at Brigham Young University, and the IBM 7040/7094 at the University of Washington. Several of the programs used are in general use, however. The program used to test for "goodness of fit" of the transformation to equation [6] was STAT 03 from the Brigham Young University Computer Research Center. The generalized sorting routine from this center was used on all tape sorts. The program BMD02V of the BMD series [Dixon (1965)] was used for all the factorial analyses included in the study. This program was run on the IBM 7094 at the University of Washington. BMD09S was used to convert to logarithms and antilogs to transform the regression coefficients into the form in which the factors are presented by most workers [see the introduction] for comparisons. The regression program giving Figure 27 as output is also quite specialized, but very simple, and therefore was also not included.

VARIABLES AFFECTING THE LACTATION ESTIMATION CURVE IN WESTERN DAIRY CATTLE

by

VERN JAY CRANDALL

B. A., Brigham Young University, 1963

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1966

Approved by:

Stanley Wearden

ABSTRACT

A study was made on records of 18,541 cows from North Dakota, Iowa, Nebraska, Montana, Idaho, Wyoming, New Mexico, Arizona, Colorado, Utah, and Nevada, records processed on a monthly basis by the DHI Computing Service, Provo, Utah. Eight variables were studied to determine their effect on the shape of the lactation estimation curve, i.e., the curve representing the factors used to estimate total 305-day from cumulative production to date. The variables, i.e., geographic area, per cent fat, lactation number, season, previous days dry, days open, level of milk production, and level of fat production, were broken into five analyses and studied in terms of their effect on both the milk and fat lactation estimation curves.

The ratios of total to part milk and fat production were computed for every stage of the lactation in terms of days in milk, where days in milk varied from one to 305 days, for each cow in the study. The date for each cow were transformed, using a log-log transformation, to give one regression coefficient characterizing the entire lactation estimation curve. In the five factorial analyses for milk and fat, the common regression coefficient for each cell was considered the observation to be analyzed, with the exception of the analyses for level of production, where the individual records were considered the observations.

The analysis of variance, the analysis of components of variance, the analysis of means, and the analysis of interactions were used in the study.

The analysis of variance for the five analyses indicated that Per Cent Fat, Lactation Number, Season, and Level of Milk and Fat Production were significant for both milk and fat coefficients. Days Open and Days Previous Dry were found to be not significant. Per Cent Fat was found to interact with Season in almost every case, in one case, Lactation Number interacted with Geographic Area, and in two instances, Per Cent Fat interacted with Geographic Area.

The analysis of the components of variance indicated that in all cases, Lactation Number contributed the largest percentage of the variance. Per Cent Fat and Season contributed more to the total variance for milk coefficients than for fat coefficients.

The multiple comparisons of means using Tukey's hsd test indicated a different structure for the milk coefficients than for the fat coefficients, although these results, in terms of which levels could be combined for each of the variables, could not always be taken as representative because of the presence of significant interaction.

The analysis of the interaction effects found to be significant indicated that Per Cent Fat by Season interactions were quite consistent and important relationships both with respect to milk and to fat. It was suggested that both Breed

and level of Average Per Cent Fat played roles in determining both the interaction and the per cent fat effect, because of the definite jumps in the graphs of the data for the high fat level cows.

Because of the effects of the interaction in most of the relationships, the factors were given in terms of the cell common regression coefficients for each of the original cells of the analysis, with the idea that further research needed to be done to separate and study the Breed by Per Cent Fat by Season effects.